

UNIVERSITÉ DU QUÉBEC

MÉMOIRE PRÉSENTÉ À  
L'UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

COMME EXIGENCE PARTIELLE  
DE LA MAÎTRISE EN BIOPHYSIQUE ET BIOLOGIE CELLULAIRES

PAR  
CATHERINE ST-LOUIS

VÉRIFICATION DE L'IMPACT DU NOMBRE D'HÉLICES  
SUR LA STABILITÉ D'UNE PROTÉINE TOUTE ALPHA

JUILLET 2007

Université du Québec à Trois-Rivières

Service de la bibliothèque

Avertissement

L'auteur de ce mémoire ou de cette thèse a autorisé l'Université du Québec à Trois-Rivières à diffuser, à des fins non lucratives, une copie de son mémoire ou de sa thèse.

Cette diffusion n'entraîne pas une renonciation de la part de l'auteur à ses droits de propriété intellectuelle, incluant le droit d'auteur, sur ce mémoire ou cette thèse. Notamment, la reproduction ou la publication de la totalité ou d'une partie importante de ce mémoire ou de cette thèse requiert son autorisation.

## REMERCIEMENTS

Tout d'abord, je tiens à remercier monsieur Marc Beauregard, mon directeur de recherche. Sa direction toujours pertinente, son encadrement et son encouragement lors de moments difficiles m'ont permis de mener à bien ce projet de manière progressive et indépendante. De plus, le climat de travail exceptionnel, parsemé de moments de délicates folies et de ridicule reste pour moi des souvenirs précieux que je garderai à jamais.

Merci Marc!

Il me serait impossible de continuer sans remercier mes collègues de laboratoire depuis le début; Maxime Sasseville, Sébastien Claveau, Cindy Lavallée, Mylène Gagnon, Vicky Houle, Étienne Dubé et Fatma Meddeb, ainsi que Daniel Leblanc qui a participé au projet en tant que stagiaire estival. Leurs appuis techniques et leurs apports scientifiques m'ont grandement aidé lors de cette entreprise excitante. Il en est de même pour le corps professoral et les technicien(ne)s de laboratoire. J'aimerais évidemment remercier aussi tous les membres de ma famille. Sans tout comprendre ce que je leur ai dit sur le sujet durant ces quelques années, ces derniers ont toujours su trouver les bons mots pour m'aider, que ce soit au niveau de l'encouragement ou même au niveau scientifique. Tous ces gens ont, directement ou indirectement, fait partie de mon équipe technique et morale tout au long de mes travaux.

Encore une fois, merci.

## RÉSUMÉ

La protéine MB-1 est une protéine synthétique conçue pour combler les manques nutritionnels chez la vache laitière tout en permettant d'avancer les connaissances sur les structures à carrés d'hélices. Plusieurs projets d'amélioration de la protéine ont été effectués. Certains d'entre eux ont vu l'apparition de mutants stables à nombre d'hélices différents de quatre, soient des mutants codant pour deux ou six hélices. Le but du présent travail est de vérifier la stabilité conformationnelle et la résistance aux protéases de ces mutants. De plus, l'impact du remplacement d'un site connu comme étant sensible aux protéases d'*Escherichia coli* chez la protéine originale sera étudié sur ces mutants. La stratégie élaborée pour ce projet ainsi que les étapes de biologie moléculaire, d'expression, de purification et de caractérisation des mutants seront délimitées. Il semblerait que les protéines à deux hélices sont totalement instables. Il en est de même pour le mutant codant pour six hélices. L'état différent des mutants (en fusion avec une autre protéine ou non) est la raison la plus probable expliquant la disparité de résultats entre les études antérieures et la présente. En ce qui concerne la modification du site sensible aux protéases, les résultats sont variés. Ce changement augmente l'accumulation des mutants à quatre et six hélices dans le milieu bactérien, mais ce n'est pas le cas pour les mutants à deux hélices. De plus, cette augmentation de la résistance aux protéases n'est pas transposée lorsque les mutants sont en présence d'un mélange de protéases ayant un profil similaire au rumen. La conclusion finale de ce projet est que la meilleure forme parmi

celles explorées pour la séquence MB-1 est un carré à quatre hélices, tel que l'imposait son design initial.

## TABLE DES MATIÈRES

REMERCIEMENTS .....	ii
RÉSUMÉ .....	iii
TABLE DES MATIÈRES .....	v
LISTE DES TABLEAUX .....	ix
LISTE DES FIGURES .....	x
LISTE DES SYMBOLES ET ABRÉVIATIONS .....	xii
CHAPITRE 1      INTRODUCTION .....	1
1.1    Certaines applications des protéines .....	1
1.2    La nutrition chez la vache laitière .....	1
1.2.1    Sources possibles d'acides aminés .....	1
1.2.2    Acides aminés essentiels chez la vache laitière .....	2
1.3    À la rescousse de l'alimentation des animaux d'élevage – solutions possibles .....	3
1.3.1    L'ingénierie métabolique .....	4
1.3.2    L'expression hétérologue d'une protéine naturelle .....	5
1.3.3    L'expression hétérologue d'une protéine de design de novo .....	7
1.4    Le design de novo de protéines .....	8
1.5    Une nouvelle approche pour combler la demande en AAEs chez la vache laitière; la protéine MB-1 .....	9
1.5.1    Le concept initial .....	9

1.5.1.1	L'hélice $\alpha$ .....	10
1.5.1.2	Le carré d'hélices .....	10
1.5.1.3	Les détails de la séquence .....	11
1.5.2	La caractérisation de la séquence initiale .....	12
1.5.3	Une lacune majeure de la protéine .....	13
1.6	MB-1Trp; la création d'un variant stabilisé de MB-1 .....	13
1.7	Projets antérieurs menant à cette étude et hypothèse .....	14
1.8	Support de la littérature envers l'hypothèse – survol .....	15
1.9	Les objectifs de la présente étude .....	16
<b>CHAPITRE 2</b>	<b>STRATÉGIE .....</b>	<b>18</b>
2.1	Constructions génétiques menant à l'expression des divers mutants .....	18
2.1.1	Choix du vecteur d'expression .....	18
2.1.2	Mutations dirigées par « QuikChange » de Stratagene .....	22
2.1.3	Construction des gènes pour le mutant à deux hélices (« 2H ») .....	24
2.1.4	Insertions des mutations « EV » .....	27
2.1.5	Amélioration de la purification par mutations « Stop » .....	27
<b>CHAPITRE 3</b>	<b>MATÉRIELS ET MÉTHODES .....</b>	<b>29</b>
3.1	Constructions génétiques menant à l'expression des divers mutants .....	29
3.1.1	PCR mutagéniques .....	29
3.1.2	Les mutants 2H .....	29
3.1.3	Mutations « EV » .....	33
3.1.4	Mutation « Stop » .....	33

3.2	Expression et purification.....	34
3.3	Quantification des protéines.....	37
3.4	Investigation conformationnelle par fluorescence intrinsèque.....	37
3.5	Investigation conformationnelle par dichroïsme circulaire.....	37
3.6	Dénaturation thermique.....	38
3.7	Dégradation protéolytique.....	39
<b>CHAPITRE 4</b>	<b>RÉSULTATS.....</b>	<b>40</b>
4.1	Constructions génétiques.....	40
4.2	Expression et purification.....	41
4.3	Investigation conformationnelle par fluorescence intrinsèque.....	44
4.4	Investigation conformationnelle par dichroïsme circulaire.....	45
4.5	Dénaturation thermique.....	48
4.6	Dégradations protéolytiques.....	48
<b>CHAPITRE 5</b>	<b>DISCUSSION ET CONCLUSIONS.....</b>	<b>52</b>
5.1	Constructions génétiques pour à la production des mutants 2H.....	52
5.2	Expression et purification.....	52
5.2.1	Obstacles rencontrés.....	52
5.2.2	Différence majeure entre les résultats antérieurs et présents.....	55
5.3	Effet bénéfique de la mutation « EV ».....	56
5.4	Caractérisations.....	57
5.4.1	Investigation conformationnelle par fluorescence intrinsèque.....	57



5.4.2	Investigation conformationnelle par dichroïsme circulaire .....	58
5.4.3	Dénaturation thermique .....	58
5.4.4	Dégradations protéolytiques .....	59
5.5	Atteintes des objectifs .....	60
5.6	Avenues de recherche .....	61
<b>BIBLIOGRAPHIE .....</b>		<b>63</b>

## LISTE DES TABLEAUX

TABLEAU 2.1	Présentation des différents mutants de MB-1 à l'étude dans ce projet.....	19
TABLEAU 2.2	Diverses étapes de la construction des vecteurs.....	21
TABLEAU 2.3	Points isoélectriques théoriques des différents mutants à l'étude.....	28
TABLEAU 3.1	Mélange pour PCR mutagénique.....	30
TABLEAU 3.2	Conditions pour PCR mutagénique.....	30
TABLEAU 3.3	Séquences des amorces mutagéniques « Rogne ».....	31
TABLEAU 3.4	Détails des étapes de purification.....	35
TABLEAU 3.5	Détails des lectures de dichroïsme circulaire.....	38
TABLEAU 4.1	Pourcentages de structures secondaires de 4H et 4H-EV calculés via CDSSTR.....	47

## LISTE DES FIGURES

FIGURE 1.1	Heptade et carré d'hélices.....	11
FIGURE 1.2	Schéma de la protéine MB-1 originale.....	12
FIGURE 2.1	Schémas du vecteur pMAL-c2X de New England Biolabs....	21
FIGURE 2.2	Survol des étapes du protocole « QuikChange » de Stratagene.....	23
FIGURE 2.3	Emplacements prévus des sites de restriction AatII au sein du gène de MB-1TrpHis(EV) pour la construction de 2HA(EV) puis 2HB.....	25
FIGURE 2.4	Séquences pour la construction de 2HA (et par extension 2HA-EV).....	26
FIGURE 2.5	Séquences pour la construction de 2HB.....	26
FIGURE 3.1	Alignement des diverses amorces mutagéniques sur la séquence de 4H.....	32
FIGURE 3.2	Séquences des amorces mutagéniques « corrA ».....	33
FIGURE 3.3	Séquences des amorces mutagéniques « EV ».....	33
FIGURE 3.4	Séquences des amorces mutagéniques « Stop ».....	33
FIGURE 4.1	SDS-PAGE coloré au nitrate d'argent de différentes fractions recueillies au cours de la purification de 4H-EV.....	42
FIGURE 4.2	SDS-PAGE coloré au nitrate d'argent de différentes fractions recueillies au cours de la purification de 2HA.....	43
FIGURE 4.3	SDS-PAGE des différentes purifications des mutants, gel coloré au nitrate d'argent.....	44
FIGURE 4.4	Spectre d'émission fluorescente de 4H et 4H-EV lorsqu'excités à 280 nm.....	46
FIGURE 4.5	Spectre de dichroïsme circulaire de 4H et 4H-EV à 20 °C.....	47

FIGURE 4.6	Suivit de la dénaturation thermique de 4H et 4H-EV par DC à 222 nm.....	49
FIGURE 4.7	$\Delta G$ de dénaturation à 222 nm pour 4H et 4H-EV.....	50
FIGURE 4.8	Dégradations protéolytiques (Pronase E) à 39°C des différents mutants de MB-1 à 0,2 mg/mL.....	51

## LISTE DES SYMBOLES ET ABRÉVIATIONS

2HA	nouvelle forme de MB-1Trp dont la séquence équivaut aux hélices 1 et 2 – aussi valable pour le gène de la protéine
2HA-EV	2HA avec la modification du site « NL » vers « EV » – aussi valable pour le gène de la protéine
2HB	nouvelle forme de MB-1Trp dont la séquence équivaut aux hélices 3 et 4 – aussi valable pour le gène de la protéine
4H	autre appellation de MB-1Trp – aussi valable pour le gène de la protéine
4H-EV	variant stabilisé de MB-1Trp – aussi valable pour le gène de la protéine
6H	nouvelle forme de MB-1Trp dont la séquence équivaut à la suite d'hélices 1-2-1-2-3-4 – aussi valable pour le gène de la protéine
6H-EV	6H avec la modification du site « NL » vers « EV » – aussi valable pour le gène de la protéine
A	alanine
AAE	acide(s) aminé(s) essentiel(s)
ADN	acide désoxyribonucléotidique
ANSA	acide 8-anilino-1-naphtalene sulfonique
BCA	acide bicinconinique
CaCl <sub>2</sub>	chlorure de calcium
C-terminal	extrémité carboxyle de la protéine
CTPR	séquence TPR consensus
DC	dichroïsme circulaire
DHPS	dihydropicolinate synthase

dNTP	bases de l'ADN; soit... dATP (2'-deoxyadenosine 5'-triphosphate), dCTP (2'-deoxycytidine 5'-triphosphate), dGTP (2'-deoxyguanosine 5'-triphosphate) et dTTP (2'-deoxythymidine 5'-triphosphate)
DO	densité optique
E	glutamate
E.C.	« Enzyme Commission number », ou nombre de classification des enzymes
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	d'acide éthylènediaminetétracétique dihydrate
EGTA	d'acide (éthylène bis (oxyéthylènenitrolo)) tétraacétique
EV, E <sub>44</sub> V <sub>45</sub>	modification stabilisante chez MB-1TrpHis
F	phenylalanine
FPLC	Fast Protein Liquid Chromatography
G	glycine
H	histidine
I	isoleucine
IPTG	isopropylthio-β-D-galactoside
K	lysine
L	leucine
LB	Luria-Bertani
M	méthionine
MBP	« maltose-binding protein »
MBP / MB-1X / His	protéine fusion comportant MBP, mutant quelconque (X) de MB-1 et queue d'histidine
MB-1	protéine « Milk Bundle-1 »
MB-1Trp	MB-1 avec tryptophane en position 62
MB-1His	MB-1 avec une queue de six histidines
MB-1TrpHis	MB-1 avec tryptophane en position 62 et une queue de six histidines

MCS	« multi-cloning site », ou site de poly-clonage
Mq	marqueur de poids moléculaire
N	asparagine
NaCl	chlorure de sodium
NaN <sub>3</sub>	azoture de sodium
NL, N <sub>44</sub> L <sub>45</sub>	lien asparagine-leucine en position 44-45 de la protéine MB-1TrpHis et nom du projet en question
N <sub>93</sub> L <sub>94</sub>	lien asparagine-leucine en position 93-94 de la protéine MB-1TrpHis
N <sub>142</sub> L <sub>143</sub>	lien asparagine-leucine en position 142-143 des protéines 6H ou 6H-EV
PCR	« polymerisation chain reaction », ou réaction de polymérisation en chaîne
pI	point isoélectrique
pMAL / MB-1TrpHis	vecteur codant pour MB-1TrpHis
PMSF	fluorure d'alpha-phényl méthyl sulfonyle
Q	glutamine
R	arginine
RMN	résonance magnétique nucléaire
Séq.	séquence
SDS	« sodium dodecyl sulfate », ou dodécyl sulfate de sodium
SDS-PAGE	« sodium dodecyl sulfate polyacrylamide gel electrophoresis », ou électrophorèse sur gel de polyacrylamide contenant du dodécyl sulfate de sodium
T	thréonine
TE	tampon Tris-EDTA
Tm	température de fusion d'une protéine
TPR	« tetratricopeptide repeat »
Tris	2-amino-2-(hydroxyméthyl)-1,3-propanediol
V	valine
W	tryptophane

Xgal	5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside
Y	tyrosine
$\Delta G$	changement d'énergie libre



# **CHAPITRE 1**

## **INTRODUCTION**

### **1.1 Certaines applications des protéines**

Les protéines, soient elles naturelles ou modifiées, sont souvent nécessaires en industrie. Les enzymes utilisées dans les détergents, dans la production de produits pharmaceutiques ainsi que dans la transformation des aliments et du textile n'en sont que quelques exemples. À cause de leur caractère unique et de leur action spécifique, les chercheurs ont depuis longtemps flairé l'intérêt de développer les applications biotechnologiques et physiologiques des protéines.

La création de toutes nouvelles protéines synthétiques est de plus en plus courante, permettant ainsi de dicter d'avance la fonction exacte voulue, ce qui n'est pas toujours possible lors de la modification de protéines naturelles. Dans le cadre de ce travail, la protéine « Milk-bundle-1 » (MB-1) fut dessinée dans le but spécifique de pallier aux manques nutritionnels des vaches laitières qui sont coûteux à combler autrement.

### **1.2 La nutrition chez la vache laitière**

Les vaches laitières ont un système digestif très différent de celui des êtres humains. Il y a une étape de passage initial de la nourriture dans le rumen avant de se rendre à l'estomac proprement dit. Ceci modifie la quantité de certains acides aminés disponibles pour absorption par l'animal [1].

#### **1.2.1 Sources possibles d'acides aminés**

Ces ruminants obtiennent leurs acides aminés à partir de trois sources différentes. Une source externe d'acides aminés est évidemment l'alimentation.

Dans ce cas, les protéines ingérées sont dégradées et les acides aminés résultants sont assimilés par l'organisme. Les deux autres sources sont plutôt « intrinsèques ». Les micro-organismes présents naturellement dans le rumen sont une source importante d'acides aminés, tandis que le recyclage des tissus musculaires permet aussi de fournir ces molécules.

Lorsque les protéines solubles provenant de l'alimentation passent dans le rumen, une partie d'entre elles (60-70%) est directement digérée par les protéases microbiennes. Les produits de ces dégradations servent de source de carbone, d'azote et d'acides aminés pour la flore ruminale, qui les transforme en nouvelles protéines et autres molécules. Les protéines de la flore microbienne ruminale – soient-elles sécrétées ou relâchées lors de l'apoptose – sont donc une source d'acides aminés pour les ruminants. Toutefois, ce processus est accompagné d'une perte de valeur nutritionnelle de la diète de l'animal par diminution de la quantité totale d'acides aminés accessibles. Lorsque l'apport en acides aminés provenant de l'alimentation ou de la flore ruminale est insuffisant, l'organisme peut finalement recycler les protéines des tissus musculaires ou produire de nouveaux acides aminés par biosynthèse [1-2].

### **1.2.2 Acides aminés essentiels chez la vache laitière**

Il existe deux classes d'acides aminés; ceux qui sont essentiels (AAE) à la survie de l'organisme et ceux qui ne le sont pas. Les AAE proviennent uniquement de l'alimentation, car les voies métaboliques de biosynthèse de ces molécules sont absentes chez les mammifères [3]. Des vingt acides aminés naturels principalement associés aux protéines, il y en a neuf qui sont considérés comme des AAE pour les mammifères, soient; la phénylalanine (F), l'histidine (H), l'isoleucine (I), la lysine (K), la leucine (L), la méthionine (M), la thréonine (T), la valine (V) et le tryptophane (W). En plus d'intervenir dans la

croissance et le maintient physiologique de la vache laitière, les AAE sont aussi nécessaires à la production du lait.

Le contexte supplémentaire de production massive de lait chez la vache laitière domestique transforme une alimentation dite normale en carence évidente en AAE. La qualité et la quantité de lait obtenues sont dépendantes, entre autres, d'un bon approvisionnement en AAE. Les trois sources d'acides aminés délimitées précédemment sont insuffisantes pour subvenir aux besoins accrus de ces bêtes. De plus, si le manque en AAE est suffisamment grand, il est possible que l'animal subisse une perte de masse totale et devienne plus susceptible à la maladie et à l'infection [4]. Ce sont les AAE M et L, puis dans une moindre mesure T et K, qui sont les plus limitants dans l'alimentation de ces ruminants [2,4-7].

### **1.3 À la rescousse de l'alimentation des animaux d'élevage – solutions possibles**

Depuis plusieurs années, la solution des producteurs laitiers à la carence en AAE des vaches laitières est d'utiliser des capsules contenant des acides aminés libres (surtout M), des enzymes et des vitamines. Ces capsules sont suffisamment résistantes pour traverser le rumen tout en demeurant intactes. Elles peuvent ainsi libérer leur contenu en aval dans le système digestif, là où les constituants seront adéquatement digérés et absorbés par la bête. Un problème avec cette solution est le coût des produits. Selon Aventis, un des plus grands producteurs de telles capsules, le marché mondial de la supplémentation alimentaire pour animaux d'élevage se chiffre à près de 512 millions d'Euros sur une base annuelle (environ 725,6 millions en dollars canadien) [8].

L'alimentation présentement offerte aux vaches laitières, malgré le fait qu'elle provienne de sources variées et qu'elle soit supplémentée d'additifs, n'est tout simplement pas suffisante. Cette diète riche en AAE est incapable d'augmenter de façon optimale les M, T, L et K disponibles pour le métabolisme [7]. Il s'agit donc de concevoir et tenter d'autres manières de pallier au manque nutritionnel de ces ruminants. Les produits résultants doivent être peu coûteux et suffisamment efficaces pour devenir concurrentiels. La biotechnologie peut apporter une telle solution. Les prochaines sous-sections décrivent trois approches qui sont présentement explorées pour augmenter la valeur nutritive des aliments habituellement ingérés par les ruminants.

### **1.3.1 L'ingénierie métabolique**

Cette technique a pour but de modifier certaines voies métaboliques d'un organisme pour favoriser soit une déficience, soit une accumulation en métabolites. Un aspect intéressant de cette technique est qu'elle peut être appliquée de façon spécifique par rapport à la compartimentation. Il est donc possible de restreindre la modification au métabolisme à seulement certains emplacements dans l'organisme.

Dans le cadre de la production laitière, la manipulation du métabolisme par changements génétiques de plantes fourragères pourrait rendre ces aliments plus nutritifs. Ces plantes accumulent déjà des AAE, mais pas dans des ratios optimaux pour la vache laitière. On viserait donc à augmenter la quantité de M et de K présente dans une plante pour en augmenter sa valeur nutritive. Cette augmentation peut en plus être ciblée dans les feuilles de la plante, tissu qui sera ingéré par l'animal.

Plusieurs exemples d'application de l'ingénierie métabolique sont présents dans la littérature. Particulièrement, les équipes de Boerjan et de Kim [9-10] ont tenté d'augmenter l'accumulation de M dans différents organismes.

Inopportunément, cette augmentation fut accompagnée de changements morphologiques et de fortes odeurs. D'autres projets n'ont pas permis l'augmentation attendue du taux de M [11-12]. La situation est tout autre pour la K. Notamment, l'équipe de Falco a pu modifier le soya pour permettre une accumulation de lysine deux fois plus grande que celle de l'organisme original. La modification au métabolisme est permise par le remplacement de la dihydropicolinate synthase (DHPS) originale du soya par l'enzyme bactérienne. Cette enzyme fait partie de la voie de synthèse de la K. L'enzyme végétale s'inactive lorsque le taux de K intracellulaire atteint un certain niveau. L'enzyme bactérienne, par contre, est insensible à cette rétro-inhibition, permettant une plus grande accumulation finale de K [13].

Les augmentations les plus significatives en M et K ont souvent été accompagnées de troubles de germination et de croissance. Ces phénomènes ne sont que partiellement compris à ce jour, mais il reste que la modification du métabolisme de ces AAE cause des réactions imprévisibles chez la plante. Parallèlement, la disponibilité, la correcte absorption et le transport du soufre sont importants pour le métabolisme de la M. Une carence en soufre peut donc limiter l'accumulation de M dans les plantes [14].

### **1.3.2 L'expression hétérologue d'une protéine naturelle**

Au lieu de manipuler le métabolisme des AAE, il est possible d'imposer la production d'une protéine naturelle à haut taux en AAE provenant d'un organisme initial dans un nouvel hôte. La polymérisation de ces AAE sous forme de protéine empêche leur dégradation, permet l'entreposage pour usage ultérieur et empêche une rétro-inhibition de synthèse d'AAE endogène (le cas échéant).

Une fois de plus, dans le cas de la vache laitière, ce sont M et K qui sont plutôt ciblés. Cependant, le nombre de protéines naturelles qui contiennent un

fort taux de ces AAE est limité. Notamment, l'albumine 2S de la noix du Brésil, celle du tournesol et les zéines de maïs ont un taux significativement plus élevé en M [14-17]. L'intégration d'une de ces protéines chez un hôte pertinent viendrait donc aider à combler la carence en M chez la vache. Cette stratégie peut évidemment s'appliquer aux plantes fourragères et aux céréales composant la nourriture de la vache, dont le colza, le maïs, le soja et la luzerne.

Plusieurs groupes tentent effectivement d'améliorer la valeur nutritive de l'alimentation des mammifères par expression hétérologue. Certaines études démontrent une accumulation de M cinq à sept fois plus élevée que la normale, alors que d'autres ne rapportent aucune différence. Il est possible que le soufre, composante de la M, ait été limitant pour ces dernières études. Lorsque tel est le cas, l'expression de la protéine exogène (riche en M) diminue la quantité totale des protéines riches en M initialement retrouvée chez l'organisme [18-19]. En ce qui concerne les essais pour augmenter la quantité de K accessible, des exemples positifs sont l'expression d'un homologue de l'inhibiteur de la chymotrypsine-2 et de la protéine S-VSP [20-21].

Cette approche a aussi été mise en œuvre avec les bactéries de la flore ruminale. La flore ruminale est constituée en grande partie par les espèces *Butirivibrio fibrisolvens*, *Prevotella ruminicola* et *Streptococcus bovis* [22]. Une partie de ces populations est transportée plus loin dans le système digestif de la vache avec la nourriture ingérée. Ces organismes servent donc de « capsules naturelles », protégeant leur contenu en AAE de la dégradation par le rumen. Il serait possible de modifier certaines de ces bactéries pour qu'elles produisent une ou plusieurs nouvelles protéines nutritives naturellement retrouvées chez un autre organisme. Ces bactéries seraient donc inoculées dans le rumen d'où elles proviennent et une partie d'entre elles devient ensuite les « capsules » qui sont digérées, libérant les AAE pour assimilation par le ruminant [23].

Il est possible de combiner les approches d'ingénierie métabolique et d'expression hétérologue. On rapporte une expérience d'ingénierie métabolique sur le maïs où l'on obtient une accumulation de la protéine BHL8 de l'orge (riche en K), et ce, jusqu'à un rendement de trois à six pourcents des protéines totales au niveau de la graine. Lorsque cette même souche de maïs fut aussi modifiée pour permettre l'expression hétérologue de l'enzyme DHPS en parallèle à la BHL8, la quantité de K totale passe de 0,2% du poids sec de la graine à 0,7% [24].

### **1.3.3 L'expression hétérologue d'une protéine de design *de novo***

Cette technique est en quelque sorte une extension de celle présentée dans la section précédente. Il s'agit, dans ce cas, non pas d'exprimer une protéine naturelle, mais plutôt une protéine synthétique chez l'hôte ciblé. Cette différence confère un avantage certain à la technique – la flexibilité. Au lieu de modifier la quantité d'un seul AAE à la fois avec une protéine naturelle, il est possible d'imposer un biais important de plusieurs AAE à la fois dans une toute nouvelle protéine. La conception de la protéine synthétique permet d'imposer un biais contrôlé en AAE quant à leur identité et leur proportion, dépendamment des carences à combler.

La majorité des protéines de design *de novo* retrouvées dans la littérature ont toutes deux aspects en commun. On rencontre presque constamment soit des constructions simples, soit le manque presque total de méthionine dans la séquence [25-30]. Tous les scientifiques évitent l'insertion de M dans leur design, vu la faible tendance de cet acide aminé à spécifier un repliement unique et sa faible énergie de transfert au solvant. Cet acide aminé présente une chaîne hydrophobe qui contient un atome de soufre, ce qui augmente la polarité du résidu et rend son enfouissement dans le cœur de la protéine difficile [31]. Il est possible que ces caractères de la M expliquent le faible rendement

d'expression *in vivo* (chez les bactéries) de protéines synthétiques à haut taux de cet AAE à ce jour [32-33].

Une protéine synthétique a cependant été dessinée avec un contenu élevé en M. C'est par des études comparatives de séquences dont le repliement se ressemble que des chercheurs ont établi une approche efficace pour le design de cette protéine [34].

#### **1.4 Le design *de novo* de protéines**

L'approche habituelle dans le design de protéines est de d'abord choisir un repliement final pour la protéine. Le choix peut s'inspirer de la forme de protéines naturelles ou être complètement dicté selon les idées des créateurs. Souvent, les chercheurs choisissent des formes qui présentent des structures secondaires uniques, des séquences hautement répétitives et une haute symétrie. Quelques exemples de repliements rapportés dans la littérature sont les doigts de zinc ainsi que les carrés et torsades d'hélices [26-30,35-40].

Une des principales lacunes des premiers projets de design *de novo* est sans doute le comportement peu natif des protéines générées. Pour qu'une protéine soit dite en état natif, elle doit se stabiliser via des interactions hydrophobes, l'empaquetage spécifique des chaînes latérales des acides aminés la constituant, des réseaux de ponts hydrogènes et dans certains cas des interactions électrostatiques. De plus, l'état natif est caractérisé par un état oligomérique et conformationnel unique, sujet à une dénaturation thermique ou chimique coopérative [41].

Souvent, les protéines de design *de novo* se comportent plutôt comme des globules fondus. Les protéines dans cet état peuvent être considérées comme des intermédiaires entre les états natif et dénaturé. Ces protéines ne sont que partiellement repliées; on y retrouve les diverses structures secondaires



formées, mais la structure tertiaire finale y est mal définie. La flexibilité du cœur hydrophobe permet la formation de plusieurs différents conformères. Ce manque d'homogénéité rend difficile l'étude de ces protéines. Notamment, on retrouve des problèmes d'agrégation et de dégradation protéolytique accrue [42-44].

Il existe quelques moyens pour tenter de rapprocher le globule fondu d'un état plus natif, et du même coup améliorer l'atteinte de la structure cible. D'une part, l'utilisation d'algorithmes de prédiction de repliement permet de vérifier *in silico* l'effet de différentes altérations à la séquence sur l'ensemble de la protéine, soit son cœur hydrophobe, sa surface exposée au solvant et ses régions interfaciales [39-45]. D'autre part, il est possible d'avoir recours à l'évolution dirigée. Cette technique est basée sur l'apparition d'erreurs au sein du gène codant pour la protéine et permet l'accélération du processus de sélection naturelle. Elle peut être appliquée à l'ensemble du gène (aléatoire) ou cibler de positions précises (semi-rationnelle). Les études de Chu [46] et Wang [47] sont des exemples d'évolution dirigée.

## **1.5 Une nouvelle approche pour combler la demande en AAEs chez la vache laitière; la protéine MB-1**

### **1.5.1 Le concept initial**

Dessinée pour pallier à la carence en AAE de la diète des vaches laitières, une protéine devrait contenir de forts taux de M, T, L et K. Évidemment, la protéine finale devait aussi être en mesure de résister au milieu ruminal. En plus, et ce plutôt par souci de simplicité pour faciliter la purification et les caractérisations subséquentes, la protéine devait être soluble. Grâce aux principes de repliement des protéines établis à l'époque de sa création, il fut possible de générer la séquence primaire de MB-1 et atteindre la structure tertiaire finale choisie [34].

Cette structure, choisie selon les études contemporaines avancées sur ses principes de repliement [48], est un carré d'hélices [34]. Un aspect intéressant des AAE à inclure dans la séquence aida beaucoup dans le choix de cette structure. En effet, il est reconnu que M, L et K sont d'assez bons inducteurs d'hélices [49].

#### **1.5.1.1 L'hélice $\alpha$**

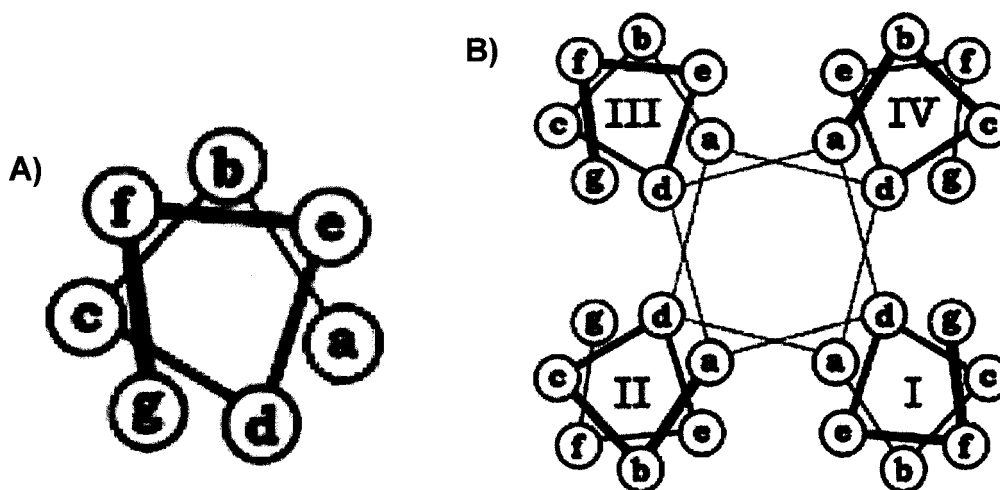
On appelle structure secondaire l'arrangement spécifique du squelette de la protéine dans l'espace. On retrouve majoritairement deux types de structures secondaires périodiques dans les protéines, soient les feuillets  $\beta$  et les hélices  $\alpha$ . L'hélice  $\alpha$  est un arrangement spiralé d'acides aminés comportant sept positions (heptade, délimitées de « a » à « g ») équivalant à un peu moins que deux tours complets d'hélice (figure 1.1A) [50]. Les acides aminés aux positions « a » et « d » sont la fraction hydrophobe de l'hélice, alors que les acides aminés aux positions « b », « c » et « f » en sont la partie hydrophile. En ce qui a trait aux acides aminés aux positions « e » et « g », ils peuvent être plus ou moins exposés à l'eau.

#### **1.5.1.2 Le carré d'hélices**

On appelle structure tertiaire l'arrangement spécifique des structures secondaires d'une protéine de manière à obtenir un cœur non polaire. On connaît bien et depuis longtemps la structure tertiaire connue sous le nom de carré d'hélices (figure 1.1B) [50]. C'est une structure à quatre hélices  $\alpha$  contiguës reliées par des connecteurs apériodiques.

La formation de cette structure est induite par les acides aminés en positions d'heptade « a » et « d ». On se rappelle que ces positions sont occupées par des acides aminés hydrophobes. Il est connu que la plus grande force poussant le repliement des protéines solubles globulaires est l'effet hydrophobe [51]. Ce phénomène a lieu car les molécules non polaires sont

repoussées les unes vers les autres par un milieu polaire, comme l'eau, qui est incapable d'incorporer de telles molécules dans son réseau de ponts hydrogènes. Dans le cas des protéines, cela implique le rapprochement des acides aminés hydrophobes en son cœur. Grâce à cette force, les hélices du carré se rapprocheront l'une de l'autre par leurs surfaces hydrophobes.



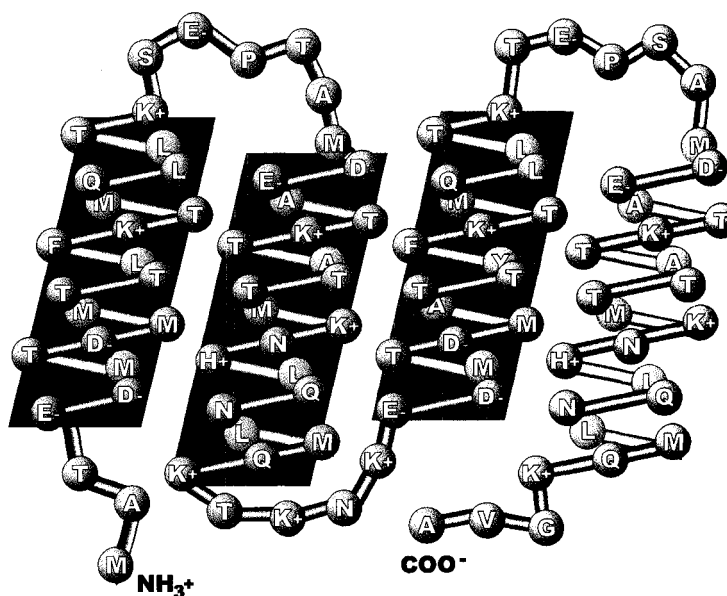
**FIGURE 1.1** Heptade et carré d'hélices. **A)** Vue schématique d'une heptade (vue de haut). **B)** Exemple d'une protéine sous forme de carré d'hélices (vue de haut) [50].

### 1.5.1.3 Les détails de la séquence

MB-1 est une protéine totalement synthétique. Les scientifiques l'ayant conçue ont donc utilisé le gabarit de carré d'hélices tel que défini par une étude de quatre protéines naturelles pour la conception de la protéine. La structure primaire (séquence d'acides aminés) de MB-1 fut ainsi construite.

La polarité, la charge et le volume des différents acides aminés furent pris en considération pour combler les différentes positions des heptades de la protéine. La contrainte de haute valeur nutritive imposée à la protéine fut

respectée; des 100 acides aminés de la séquence finale, 57 d'entre eux sont des M, T, L et K (figure 1.2). Une fois la séquence finale en acides aminés décidée, le code génétique de la bactérie *Escherichia coli* (*E. coli*) fut utilisé pour créer un gène optimisé pour la production de la protéine dans cet hôte [34].



**FIGURE 1.2** Schéma de la protéine MB-1 originale. On dénote la position prédite des hélices (rectangles de couleur). Les acides aminés en avant-plan sont ceux qui sont exposés à la surface de la protéine, et ceux en arrière-plan font partie du cœur hydrophobe.

### 1.5.2 La caractérisation de la séquence initiale

Avant toute chose, il était pertinent de vérifier la valeur nutritive de MB-1 ainsi que sa taille. Tel que prévu, la protéine contenait 57% des AAE M, T, L et K, puis présentait une masse moléculaire d'environ 11 kDa. Quant à la structure elle-même, elle semblait conforme aux attentes pour une structure de carré d'hélices. Des analyses de dichroïsme circulaire (DC) confirmaient une grande quantité de structures hélicales (environ 50%), puis des études de

fluorescence intrinsèque ont démontré que la structure tertiaire était suffisamment bien formée pour cacher la tyrosine (Y) en position 62 [34,52].

D'autres expériences ont aussi illustré le caractère quasi natif de MB-1. La faible affinité pour l'acide 8-anilino-1-naphtalene sulfonique (ANSA) suggérait un cœur hydrophobe suffisamment compact, semblable à certaines protéines natives [34,35]. De plus, la dénaturation chimique de MB-1 par l'urée était réversible [34].

### **1.5.3 Une lacune majeure de la protéine**

Malgré le fait que MB-1 se comporte de façon plutôt native, elle présenterait une grande susceptibilité aux conditions ruminales. Deux aspects clés permettent de prédire ceci. On note que la température de fusion ( $T_m$ ) de la protéine est celle de la température présente dans le rumen de la vache laitière, soit 39°C. La moitié des protéines MB-1 est donc sous forme dénaturée dans le rumen.

Cette première faiblesse mène à une deuxième, soit une dégradation protéolytique accrue. En effet, lorsqu'on incube la protéine à 39°C en présence de protéases à spectre large imitant les protéases ruminales (le mélange Pronase E), MB-1 est facilement dégradée [52,54]. Il en résulte que la protéine MB-1 n'est pas apte à être utilisée, sous sa forme originale, comme supplément alimentaire chez la vache laitière.

## **1.6 MB-1Trp; la création d'un variant stabilisé de MB-1**

Une version améliorée de MB-1, MB-1Trp (nommé 4H à partir de ce point), fut développée en 2000. La stratégie consistait à augmenter la résistance aux protéases par une hausse de la stabilité conformationnelle de la protéine. Pour ce faire, la sonde spectroscopique interne de MB-1 – la Y de la troisième hélice,

en position 62 (voir figure 1.2) – fut remplacée par un W plus hydrophobe et plus volumineux. La Y présente un groupement hydroxyle non satisfait parmi les alanines l'entourant. Le W, par contre, ne présente pas de groupement ayant besoin de stabilisation, puis remplirait mieux la poche formée par les alanines (A) au cœur de la protéine ( $187 \text{ \AA}^3$  pour Y versus  $217 \text{ \AA}^3$  pour W) [41]. L'hypothèse mise de l'avant par les auteurs était que cette modification aiderait à la compaction de la protéine en apportant un plus grand effet hydrophobe au cœur de la protéine. C'est bien ce qui s'est produit. Le  $T_m$  de la protéine fut augmenté de  $39^\circ\text{C}$  à  $55^\circ\text{C}$ , puis la résistance aux protéases est devenue voisine de celle du cytochrome c (protéine globulaire de taille similaire à MB-1) [55].

### 1.7 Projets antérieurs menant à cette étude et hypothèse

Les résultats de plusieurs projets antérieurs ont mené au présent projet. Tout d'abord, un projet a permis la construction d'une MB-1 avec une queue de six H en carboxy-terminal (C-terminal) de la protéine (« MB-1His »). On tentait ainsi d'augmenter le degré final de pureté de la protéine en changeant la dernière chromatographie, une échangeuse d'ion, par une chromatographie à affinité pour un métal immobilisé (le nickel dans ce cas). Il fut donc possible de purifier la nouvelle protéine « MB-1His », et ce, à une plus grande pureté que la MB-1 originale, tel qu'attendu [56]. Les différents mutants de MB-1 qui ont suivi lors des divers projets étaient aussi munis de cette queue d'histidine (MB-1TrpHis, etc...). Ce projet est important par le fait qu'il s'est justement greffé à tous les autres le suivant.

Un autre projet ayant une importance plus grande dans ce présent projet est nommé « Alanines de cœur ». Il s'agissait d'une tentative d'amélioration de la compaction et la stabilité de la protéine par l'augmentation du volume du cœur de la protéine. Pour ce faire, les A se retrouvant dans le cœur de la protéine furent remplacés par des acides aminés hydrophobes plus volumineux, de façon combinatoire. Une banque de 7776 mutants différents de MB-1TrpHis

fut générée, puis exposée à des conditions sévères dans le but de sélectionner les protéines les plus stables. Dans ce cas, le critère de sélection était la résistance aux protéases (Pronase E, un mélange de protéases similaires permettant de mimer le rumen des vaches) et à la chaleur. Les essais ont permis de découvrir plusieurs mutants plus stables que la protéine d'origine. Cependant, contre toutes attentes, certains d'entre eux n'adopteraient pas la forme de carrés à quatre hélices alpha. La nouvelle structure la plus souvent rencontrée lors de ce projet semble ne comporter que deux hélices, soient les deux premières (1 et 2) ou les deux dernières (3 et 4) de MB-1TrpHis [57].

Le deuxième projet préalable à mes travaux, nommé « NL », visait à augmenter la résistance de MB-1TrpHis aux protéases d'*E. coli* dans le but d'augmenter l'accumulation de la protéine dans la bactérie. Les expériences de mutagenèse dirigée sur un site asparagine-leucine (NL), identifié comme étant sensible aux protéases d'*E. coli*, ont permis de produire une banque de 20 mutants. Cette banque fut exposée aux protéases d'*E. coli* pour discerner les plus stables. L'étude démontre que le remplacement du site N<sub>44</sub>L<sub>45</sub> vers la combinaison glutamate-valine (EV) est favorable et augmente la résistance à la protéolyse. De plus, ici encore un autre mutant inattendu fut identifié. Ce dernier a un poids moléculaire équivalant à six hélices de MB-1TrpHis. Cette version alourdie démontre une stabilité apparente accrue lorsque comparée à celle de la protéine d'origine. Cette forme est une protéine dont la séquence équivaut à six hélices retrouvées dans MB-1, soit la suite d'hélices 1-2-1-2-3-4, chez qui on ne note aucune modification du site de protéolyse (les sites « NL » des hélices « 2 » ne sont pas modifiés) [58].

### **1.8 Support de la littérature envers l'hypothèse - survol**

Les règles qui régissent le bon repliement et la stabilité des protéines en général n'étant pas encore tout à fait élucidées, le dessin de nouvelles protéines stables est difficile. Plusieurs études ont déjà été effectuées sur des peptides et

protéines de design *de novo*, qui sont totalement synthétiques ou dérivés de protéines naturelles, qui ne sont composés que d'hélices  $\alpha$  et de connecteurs.

Certaines de ces études se concentrent sur les carrés d'hélices pour comprendre leur repliement, leurs fonctions inhérentes et la relation entre ces derniers [59]. Il y a aussi des recherches qui tentent de déterminer les forces exactes (incluant l'emplacement d'acides aminés spécifiques) qui dirigent la formation de ressorts superenroulés (coiled-coils), des structures à deux hélices [60,61]. Des protéines *de novo* constituées de six [62] et même neuf [63] hélices ont aussi déjà été produites et caractérisées.

Parmi toutes les études pertinentes, celles d'une équipe en particulier ressortent du lot. Main et ses collaborateurs étudient des protéines synthétiques comportant différents nombres de structures « tetratricopeptide repeat », ou TPR. Ces structures sont des arrangements hélice-connecteur-hélice de 34 acides aminés de longueur se retrouvant dans plus de 300 protéines naturelles différentes et s'y répétant entre trois et seize fois. Cette équipe a d'abord décidé d'une séquence TPR consensus (« CTPR »), puis ont répété ce consensus différents nombres de fois pour obtenir des gènes codant pour différents nombres de TRP (« CTPR-1 » pour le TPR simple, « CTPR-2 » pour deux unités, etc...) Ce qui découle de ces études [64-66] est que plus on augmente le nombre de répétition de la séquence consensus, plus la protéine produite est stable. Le taux de plus en plus lent de dépliement des protéines lorsqu'on augmente le nombre de répétition est la raison que Main offre pour expliquer la stabilisation.

### **1.9 Les objectifs de la présente étude**

L'hypothèse que j'avance dans le cadre de ce travail est que 4H serait plus stable si elle adoptait une autre structure qu'un carré d'hélices, soit en tant que structure à deux ou à six hélices. Il est possible que la diminution de la surface



hydrophobe soit bénéfique et expliquerait en partie la stabilisation des mutants à deux hélices. En contrepartie, une meilleure compaction du cœur de la protéine ou même un changement majeur de conformation peut expliquer la stabilisation du mutant à six hélices. Ces points seront vérifiés lors du présent projet. En parallèle à cette hypothèse première, il serait intéressant de vérifier l'impact de la mutation « EV », déjà reconnue comme étant stabilisante chez MB-1TrpHis [57], sur ces nouvelles structures lorsque possible.

Les étapes suivantes seront réalisées pour mener à bien ce travail :

1. Faire les manipulations génétiques qui permettront la production des différents mutants de MB-1 à l'étude par *E. coli*,
2. Produire, purifier et caractériser ces mutants,
3. Comparer les propriétés des mutants à 2 et 6 hélices à celles de la protéine originale à 4 hélices (4H),
4. Comparer les propriétés des mutants comportant la modification « EV » (à 2, 4 et 6 hélices) à celles de leur contrepartie originale.

## CHAPITRE 2

### STRATÉGIE

#### 2.1 Constructions génétiques menant à l'expression des divers mutants

Sept protéines seront à l'étude durant ce projet. Ces protéines se divisent en trois familles selon le nombre d'hélices de MB-1 qui sont produites. Elles sont présentées au tableau 2.1.

Pour ce projet, certains vecteurs ont dû être complètement construits alors que d'autres nécessitaient quelques modifications. Cette section présente la stratégie utilisée lors de la préparation des vecteurs. Un survol des différentes étapes pour les divers mutants est présenté au tableau 2.2, tandis que les détails expérimentaux reliés à ces étapes seront divulgués dans la section « Matériaux et méthodes ».

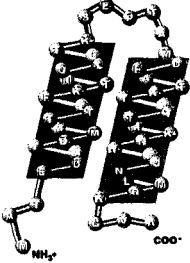
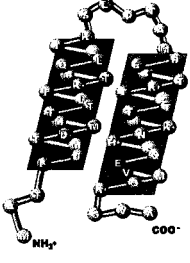
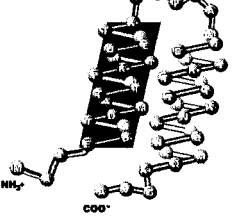
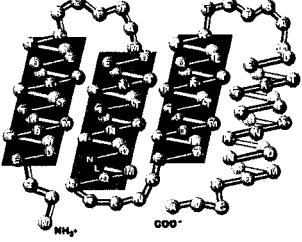
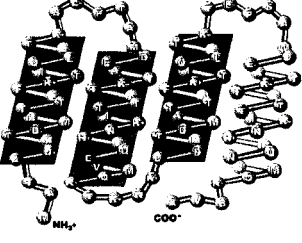
##### 2.1.1 Choix du vecteur d'expression

Le vecteur pMAL-c2 de New England Biolabs (figure 2.1) [67] fut choisi pour diverses raisons. Premièrement, la présence du gène *Amp<sup>r</sup>* confère une résistance à l'antibiotique ampicilline à l'hôte bactérien (*E. coli*), permettant la croissance sélective de cet organisme transformé avec le vecteur sur milieu contenant le dit antibiotique. Deuxièmement, la séquence pBR322 ori permet la réplication du vecteur au sein d'*E. coli*. La séquence « M13 ori » permet la production d'ADN simple brin, mais ne sera pas utilisé dans ce projet. Troisièmement, le *rrnB* terminator donne le signal de fin de transcription. Finalement, les séquences importantes pour l'expression des protéines de ce projet sont *lacI<sup>q</sup>*, *P<sub>tac</sub>*, *malE*, polylinker et *lacZα*.

La séquence « polylinker » ou « multi cloning site » (MCS) présente plusieurs sites de restriction exclusifs à cette section. Une double digestion du

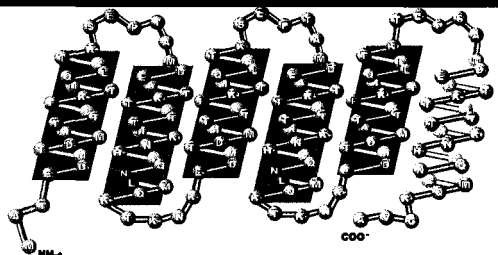
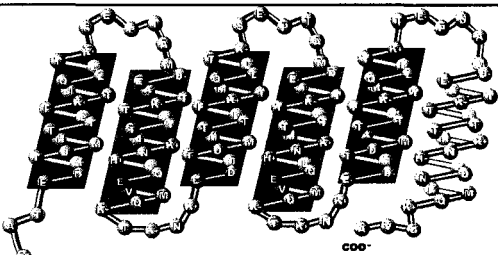
TABLEAU 2.1

Présentation des différents mutants de MB-1 à l'étude dans ce projet.

Famille	Protéine (mutant)	Modèle
À deux hélices	2HA	
	2HA-EV	
	2HB	
À quatre hélices	4H	
	4H-EV	

**TABLEAU 2.1 (suite)**

Présentation des différents mutants de MB-1 à l'étude dans ce projet.

Famille	Protéine (mutant)	Modèle
À six hélices	6H	
	6H-EV	

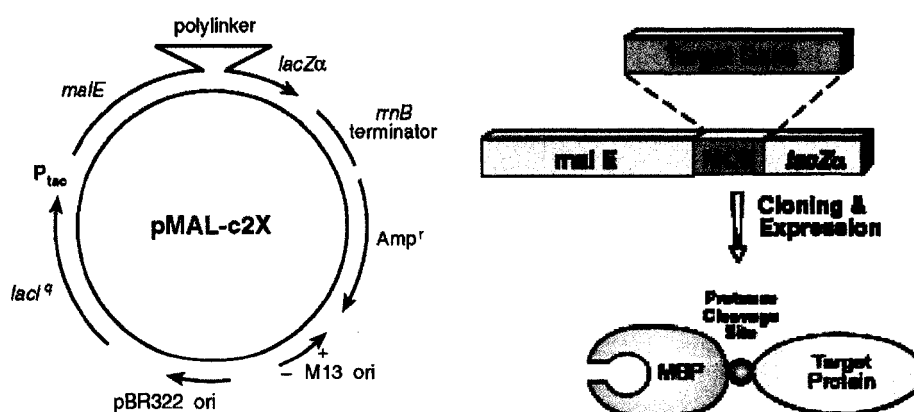
vecteur avec des enzymes coupant au sein de cette séquence permet le clonage directionnel des gènes à l'étude.

Dans le cas présent, les différents gènes « MB-1X His » sont déjà clonés dans pMAL-c2 via les sites de restriction *EcoRI* et *HindIII*. Il est possible de vérifier si nous avons bien insertion du gène grâce à la séquence *lacZα*. L'insertion du gène d'intérêt dans le MSC interrompt le gène *lacZα* et l'expression de l'enzyme  $\beta$ -galactosidase est impossible. Ceci permet un criblage au 5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside (Xgal) sur pétri des bactéries comportant les plasmides modifiés. Les colonies comportant les plasmides non modifiés sont colorées bleu grâce à l'activité du fragment  $\alpha$  de la  $\beta$ -galactosidase alors que l'insertion d'un gène au sein du MSC du plasmide rend impossible l'activité de cette enzyme puis rend blanches les colonies le transportant.

**TABEAU 2.2**

Diverses étapes de la construction des vecteurs.  
 « Mut. » = mutagenèse dirigée, recirc. = recircularisation.

Nom	ADN initial	Étape 1	Étape 2	Étape 3	Étape 4	Étape 5
2HA	MB-1 TrpHis	Mut. « 2 »	Mut. « 4 »	Digestion / recirc.	Mut. « Stop »	-
2HA-EV	4H-EV-His	Mut. « 2 »	Mut. « 4 »	Digestion / recirc.	Mut. « Stop »	-
2HB	MB-1 TrpHis	Mut. « 1 »	Mut. « 3 »	Digestion / recirc.	Mut. « corrA »	Mut. « Stop »
4H	MB-1 TrpHis	Mut. « Stop »	-	-	-	-
4H-EV	4H-EV-His	Mut. « Stop »	-	-	-	-
6H	6H-His	Mut. « Stop »	-	-	-	-
6H-EV	6H-His	Mut. « EV »	Mut. « Stop »	-	-	-



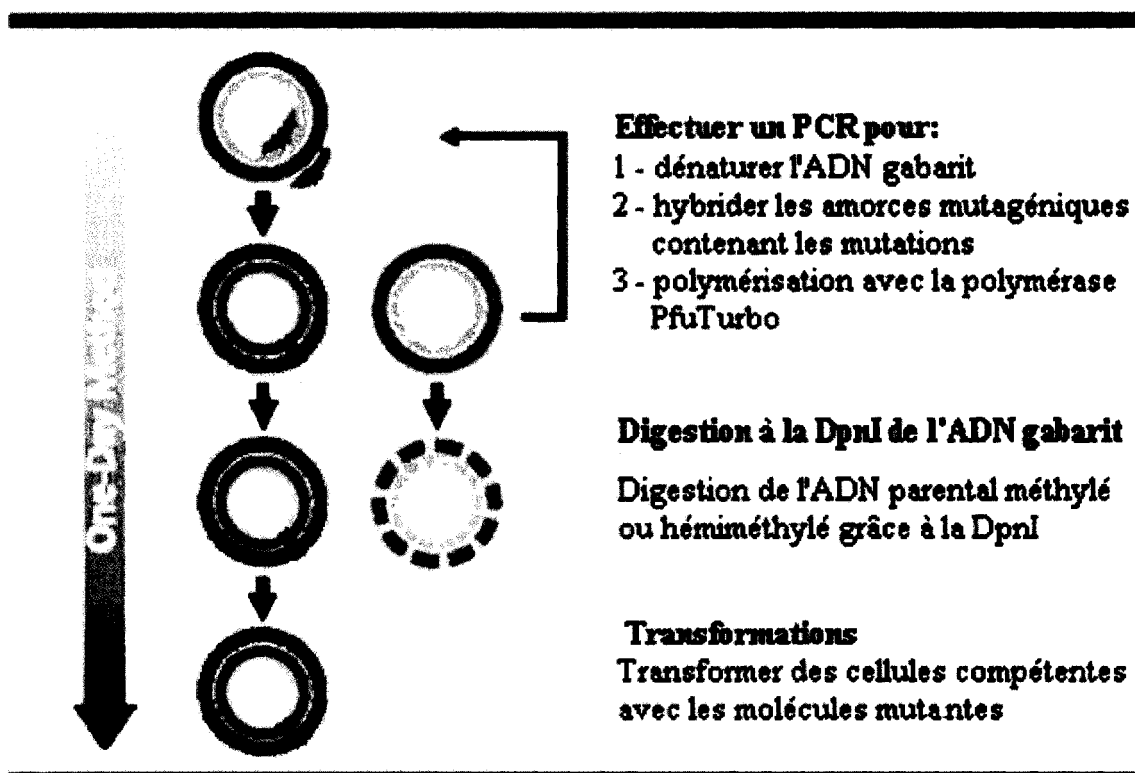
**FIGURE 2.1** Schémas du vecteur pMAL-c2X de New England Biolabs. GAUCHE, carte du vecteur; DROITE, clonage et protéine fusion ainsi exprimée (MCS = « polylinker ») [67].

L'expression des protéines du vecteur est sous le contrôle du promoteur  $P_{tac}$ . Ce promoteur est un hybride consistant de la région -35 du promoteur *trp* et la région -10 du promoteur / opérateur *lacUV5* [68]. Le promoteur lui-même est contrôlé par répression via l'allèle *lacI<sup>q</sup>* et l'expression de la protéine LacI. La répression est levée lorsqu'on ajoute de l'isopropylthio- $\beta$ -D-galactoside (IPTG) à la culture, ce qui permet la production de la protéine voulue. Ce contrôle via répression est primordial, car il empêche l'accumulation prématurée de la protéine d'intérêt qui est possiblement toxique pour l'hôte.

Dans le cas de constructions utilisant le vecteur pMAL-c2, une protéine fusion est ce qui est exprimé à l'induction. Le gène *malE* exprime la « maltose binding protein » (MBP), qui contient une séquence d'espacement en C-terminal présentant un site de clivage de la protéase spécifique facteur Xa. Le gène *malE* est aussi disposé de façon à ce qu'un gène inséré dans le MSC soit exprimé directement suite à la MBP. On exprime ainsi la protéine fusion « MBP / protéine d'intérêt ». Il est possible de purifier la protéine fusion en utilisant l'affinité que la MBP démontre pour le maltose et l'amylose. On peut ensuite facilement cliver la MBP de la protéine d'intérêt par ajout de facteur Xa. Une deuxième purification est ensuite de mise pour séparer les deux éléments et obtenir une MB-1 pure.

### 2.1.2 Mutations dirigées par « QuikChange » de Stratagene

Il a été spécifié au tableau 2.2 que la plupart des étapes pour obtenir les constructions génétiques des mutants de MB-1 sont des mutagenèses dirigées. Par souci de clarté, voici un survol du fonctionnement du protocole « QuikChange » de Stratagene (figure 2.2) [69] utilisé pour les mutagenèses dans ce projet. Ce protocole fut choisi car il permet d'effectuer les changements voulus sans avoir recours à des techniques difficiles de clonage.



**FIGURUE 2.2** Survol des étapes du protocole « QuikChange » de Stratagene [modifié de la référence 69].

Ce protocole fait appel à une paire d'amorces mutagéniques s'hybridant à la même position sur le vecteur. Une de ces amorces s'hybride au brin sens du vecteur alors que l'autre s'hybride au brin anti-sens. C'est ce détail qui permet l'amplification complète du vecteur lors d'une réaction de polymérisation en chaîne (« polymerisation chain reaction », ou PCR), obtenant ainsi un vecteur muté comportant la modification recherchée. À la fin de la réaction, on obtient une grande quantité de vecteur muté et une certaine quantité de vecteur parental résiduel, indésirable pour la suite des manipulations.

Le plasmide ayant servi de gabarit pour la création de plasmides mutants est éliminé de la réaction par l'ajout de l'enzyme de restriction *DpnI*. Cette enzyme digère les acides désoxyribonucléotidique (ADN) méthylés et/ou hémiméthylés. Les bases (dNTP, soient dATP, dCTP, dGTP et dTTP) utilisées

lors de la réaction PCR ne sont pas méthylées, alors que l'ADN parental isolé de l'hôte bactérien présente des groupements méthyles. La *DpnI* digérera donc spécifiquement les vecteurs parentaux. La transformation d'un nouvel hôte bactérien avec le vecteur muté non contaminé de vecteur parental sera ainsi possible.

### **2.1.3 Construction des gènes pour le mutant à deux hélices (« 2H »)**

Ces gènes n'étaient disponibles que sous forme de produit de PCR au début de ce projet. Plusieurs essais de clonage de ces gènes dans le vecteur pMAL-c2 ont été tentés, mais sans succès. Une stratégie plus complexe, mais plus sûre, impliquant le rognage du vecteur « pMAL / MB-1TrpHis », a donc dû être conçue pour permettre de passer de quatre à deux hélices en raccourcissant le gène initial au sein du vecteur.

Tout d'abord, il a fallu trouver une enzyme de restriction qui ne coupait pas originellement le vecteur « pMAL / MB-1TrpHis ». Aucun site de restriction *AaflI* n'était présent dans la séquence. De plus, cette enzyme est facilement disponible chez divers fournisseurs, et son coût est abordable. Ces deux aspects ont permis le choix de cette enzyme pour rogner spécifiquement le vecteur.

Deux de ces sites *AaflI* ont été insérés dans le vecteur mentionné ci-haut de manière à permettre la digestion de la région codante pour les deux hélices indésirables pour chaque mutant 2H (hélices 3-4 pour 2HA et hélices 1-2 pour 2HB). Les amorces mutagéniques « Rogne » induisent l'ajout d'un site de restriction de l'enzyme *AaflI* au sein du gène de MB-1TrpHis ou MB-1TrpE<sub>44</sub>V<sub>45</sub>His aux endroits spécifiés à la figure 2.3.

On digère donc ces vecteurs mutants pour éliminer la partie indésirable du gène original, obtenant une séquence linéaire dont les deux extrémités présen-



**A)** ATGGCTACG-----GAAGACATGACCGACATGATGACCACCCTGTTTAAAC  
**B)** ATGGCTACG-----GAAGACATGACCGACATGATGACCACCCTGTTTAAAC  
**C)** ATGGCTACGG/ACGTCGAAGACATGACCGACATGATGACCACCCTGTTTAAAC

**A)** TATGCAGCTGTTGACCAAGTCGGAACCCACGGCTATGGACGAGGCCACTAAAC  
**B)** TATGCAGCTGTTGACCAAGTCGGAACCCACGGCTATGGACGAGGCCACTAAAC  
**C)** TATGCAGCTGTTGACCAAGTCGGAACCCACGGCTATGGACGAGGCCACTAAAC

**A)** GGCTACTACAATGAAGAATCATCTTCAAACCTGATGCAGAAG-----ACTA  
**B)** GGCTACTACAATGAAGAATCATCTTCAAACCTGATGCAGAAGG/ACGTCACTA  
**C)** GGCTACTACAATGAAGAATCATCTTCAAACCTGATGCAGAAG-----ACTA

**A)** AGAACAAA-----GAAGACATGACGGACATGGCCACTACGTGGTTCAAACG  
**B)** AGAACAAA-----GAAGACATGACGGACATGGCCACTACGTGGTTCAAACG  
**C)** AGAACAAAG/ACGTCGAAGACATGACGGACATGGCCACTACGTGGTTCAAACG

**A)** ATGCAGTTGTTAACGAAGACCGAGCCCTCGGCCATGGACGAGGCCACGAAGACG  
**B)** ATGCAGTTGTTAACGAAGACCGAGCCCTCGGCCATGGACGAGGCCACGAAGACG  
**C)** ATGCAGTTGTTAACGAAGACCGAGCCCTCGGCCATGGACGAGGCCACGAAGACG

**A)** GCTACAACCATGAAAAATCATCTGCAGAACTTGATGCAAAAA-----GGCGT  
**B)** GCTACAACCATGAAAAATCATCTGCAGAACTTGATGCAAAAAAG/ACGTCGGCGT  
**C)** GCTACAACCATGAAAAATCATCTGCAGAACTTGATGCAAAAA-----GGCGT

**A)** AGCTCACCACCACCACCAC  
**B)** AGCTCACCACCACCACCAC  
**C)** AGCTCACCACCACCACCAC

**FIGURE 2.3** Emplacements prévus des sites de restriction *Aafl* au sein du gène de MB-1TrpHis(EV) en **A)** pour la construction de 2HA(EV) en **B)** puis 2HB en **C)**. Les encadrés représentent les séquences codant pour les hélices (1 = bleu, 2 = rose, 3 = vert et 4 = jaune) et la queue d'histidine (en noir). On note les sites *Aafl* en gras avec l'emplacement exact de la coupure par l'enzyme (/).

tent les mêmes bouts cohésifs. Après élimination du fragment indésirable, il s'agit ensuite de recirculariser le vecteur par ajout de ligase, obtenant ainsi le vecteur contenant le gène de MB-1 à deux hélices. Une correction finale s'impose à 2HB pour restaurer le E original remplacé par la séquence glycine-

arginine-glutamine (G-R-Q) codée par le site de restriction de *AatII*, qui est exécutée tel que décrit à la section « Matériaux et méthodes ».

Les figures 2.4 et 2.5 montrent les séquences des protéines 2H ainsi créées.

```

Séq. 4H      NLGNSGRMAT ██████████ SEPTA ██████████ T··KNK██████
Séq. 2F      -----LQNLMQKRRPKNKEDM
Séq. 4F      -----
Séq. HisStop -----
Séq. 4H      ██████████TEPSAMDEATKTATTMKNHLQNLMQK·GVAHHHHHH*SLALAVVLQ
Séq. 2F      -----
Séq. 4F      -----QNLMQKDVGVAHH-----
Séq. HisStop -----LMQK·GVA*HHHHH-----
Séq. 2HA     MAT ██████████ SEPTA ██████████ GVA*

```

**FIGURE 2.4** Séquences (séqu.) pour la construction de 2HA (et par extension 2HA-EV). Les couleurs représentent les hélices de MB-1 (1 = bleu, 2 = rose, 3 = vert et 4 = jaune). Ce qui est en gras ne fait pas partie de la protéine. La séquence qui est en rouge sera éliminée à la digestion *AatII*.

```

Séq. 4H      NNLGNSGRMAT[REDACTED]SEPTA[REDACTED]TKNK[REDACTED]
Séq. 1F      -----RMATGRQDMTDM-----
Séq. 3F      -----QKTKNKGRQD-----
Séq. corrA   ---NSGRMAT (E) .....D
Séq. HisStop -----

Séq. 4H      [REDACTED]TEPSAMDEATKTATTMKNHLQNLQMKGVAHHHHHHH*SLALAVVLQRN
Séq. 1F      -----
Séq. 3F      MTDM-----
Séq. corrA   MTDMAT-----
Séq. HisStop -----LMQKGVA*HHHHH-----

Séq.s 2HB    MAT[REDACTED]TEPSAMDEATKTATTMKNHLQNLQMKGVA*

```

**FIGURE 2.5** Séquences (séq.) pour la construction de 2HB. Les couleurs représentent les hélices de MB-1 (1 = bleu, 2 = rose, 3 = vert et 4 = jaune). Ce qui est en gras ne fait pas partie de la protéine. La séquence qui est en rouge sera éliminée à la digestion *AafII*.

#### 2.1.4 Insertions des mutations « EV »

Il a été démontré que le remplacement du site N<sub>44</sub>L<sub>45</sub> de MB-1TrpHis par EV (mutant « EV ») augmente sa résistance aux protéases d'*E. coli*, et donc aide à augmenter le rendement de purification [58]. Il est intéressant d'utiliser cette mutation stabilisante dans le cas présent, vu la production des différentes protéines (à deux, quatre ou six hélices) dans ce même organisme. De plus, on peut se demander si l'effet stabilisant est seulement au niveau de la résistance aux protéases d'*E. coli* ou s'il est aussi possiblement conformationnel. Cette possibilité sera élucidée par l'étude actuelle.

#### 2.1.5 Amélioration de la purification par mutations « Stop »

Il fut remarqué dans certains projets précédents que la purification de variants de MB-1 exprimés sous forme de fusion « MBP / MB-1X / His » pouvait être bénéfique ou désavantageuse, selon les mutants de MB-1 étudiés [56]. Cette purification faisait appel à une chromatographie d'affinité au nickel (grâce à la présence de la queue d'histidines) au lieu de la chromatographie d'échange anionique habituelle [55]. Ce dernier protocole pourrait être problématique dans le cas présent, vu la similitude des points isoélectriques (pI) des mutants 2H et de la MBP (tableau 2.3).

Lors des premières tentatives de purification des différents mutants étudiés ici avec queue histidine, la séparation des protéines s'est révélée impossible ou très fastidieuse. La décision de retourner au protocole initial de purification sans queue histidine a donc été prise, et pour ce faire, des mutations « Stop » ont été apportées aux différents vecteurs d'expression. On peut ainsi prévoir que la purification des mutants 2H par échange ionique pourrait être difficile, car les pI de ces protéines se rapprochent beaucoup de celui de la MBP.

**TABLEAU 2.3**

Points isoélectriques (pI) théoriques des différents mutants à l'étude.

Protéine	Calcul 1 <sup>a</sup>	Calcul 2 <sup>b</sup>	Calcul 3 <sup>c</sup>	pI moyen
2HA	5,86	5,88	5,56	5,77
2HA-EV	5,30	5,17	5,05	5,17
2HB	5,86	5,88	5,56	5,77
4H	7,88	7,38	6,72	7,33
4H-EV	6,41	6,53	6,08	6,34
6H	9,25	8,31	7,90	8,49
6H-EV	6,66	6,81	3,30	6,59
MBP	5,48	5,35	5,22	5,35

<sup>a</sup> Selon « CRYSPRED » de Q.E.D. Life Science Discoveries, inc [70]

<sup>b</sup> Selon « Protein Calculator » du Scripps Research Intitute [71]

<sup>c</sup> Selon « Compute pI / Mw tool » d'Expasy [72]

## CHAPITRE 3

### MATÉRIELS ET MÉTHODES

#### 3.1 Constructions génétiques menant à l'expression des divers mutants

Voici en détail les procédures expérimentales menant à l'obtention de vecteurs d'expression pour les 7 protéines à l'étude dans ce projet.

##### 3.1.1 PCR mutagéniques

Tel qu'indiqué précédemment, ces réactions seront effectuées selon le protocole « QuikChange » de Stratagene [69]. Les quantités de réactifs et les étapes pour les réactions de PCR mutagéniques sont énumérées aux tableaux 3.1 et 3.2 pour un volume final de 25  $\mu$ L.

Les réactions sont ensuite digérées grâce à l'enzyme de restriction *DpnI* (ajout de 10 U directement à la réaction PCR), ce qui élimine le plasmide parental. On peut finalement transformer la souche hôte d'*E. coli* avec le plasmide mutagénique et ainsi permettre l'amplification du vecteur. Le plasmide sera donc récupéré après extraction d'ADN (c.-à-d. « miniprep »). Chaque nouvelle mutagenèse est vérifiée par séquençage avant la poursuite des manipulations.

##### 3.1.2 Les mutants 2H

Les constructions à quatre et à six hélices (« 4H » et « 6H » respectivement) étaient déjà disponibles au début de ce projet, mais les constructions 2H ne l'étaient pas. Elles ont donc été obtenues par mutagenèse dirigée, digestion et recircularisation, suivi de correction finale à la séquence, tel que spécifié à la section « Stratégie ». Les amorces présentées au tableau 3.3 ont été utilisées pour permettre l'insertion de sites *AatII* par mutagenèse dirigée.

**TABLEAU 3.1**  
Mélange pour PCR mutagénique

Ingrédient	Quantité
Tampon Turbo 10X	2,5 µL
dNTP (10mM chacun)	0,5 µL
ADN parental	20 ng
Amorce mutagénique-F (sens)	125 ng
Amorce mutagénique-R (anti-sens)	125 ng
PfuTurbo	2.5 U
Eau stérile	Pour compléter à 25 µL

**TABLEAU 3.2**  
Conditions pour PCR mutagénique

Étape		Condition(s)
Chauffage du couvercle		105 °C
Dénaturation initiale		95 °C pour 30 s
Boucle de réaction	Nombre de cycles	18
	Segment 1	95 °C pour 30 s
	Segment 2	55 °C pour 1 min
	Segment 1	68 °C pour 15 min
Conservation finale		4 °C

**TABLEAU 3.3**  
Séquences des amorces mutagéniques « Rogne »

Rogne-	Séquence de 5' vers 3'
1F	GAAGGATGGCTACGGGACGTCAAGACATGACCGACATG
1R	CATGTCGGTCATGTCTTGACGTCCCGTAGCCATCCTTC
2F	CTTCAAAACCTGATGCAGAAGAGACGTCCTAAGAACAAAGAAGACATGAC
2R	GTCATGTCTTCTTTGTTCTTAGGACGTCTCTTCTGCATCAGGTTTTGAAG
3F	GCAGAAGACTAAGAACAAAGGACGTCAAGACATGACGGACATGG
3R	CCATGTCCGTCATGTCTTGACGTCCTTTGTTCTTAGTCTTCTGC
4F	GCAGAACTTGATGCAAAAAGACGTCCGGCGTAGCTCACCAC
4R	GTGGTGAGCTACGCCGACGTCTTTTTGCATCAAGTTCTGC

Ces sites permettent le rognage du gène original (MB-1TrpHis) et la création des différentes constructions à 2 hélices. L'utilisation des amorces 1 et 3 permet la création de la construction 2HB, alors que les amorces 2 et 4 rendent la construction du vecteur 2HA possible. Ces diverses amorces vont s'hybrider sur le gène original aux positions montrées à la figure 3.1. Les protéines ainsi obtenues sont présentées à la section « Stratégie ».

Les digestions avec l'enzyme de restriction *Aa*II sont suivies par l'ajout de 10 U d'enzyme à 600 ng de plasmide à rogner dans 20 mM 2-amino-2-(hydroxyméthyl)-1,3-propanediol (Tris)-acétate, 10 mM acétate de magnésium, 50 mM acétate de potassium et 1 mM dithiothreitol, le tout à pH 7,9. Les plasmides digérés ont ensuite été purifiés sur colonne de silice (Qiaquick PCR purification kit, Qiagen [73]) ce qui a eu pour effet d'éliminer le petit fragment de digestion indésirable. La recircularisation s'est effectuée par ligation selon le protocole « Rapid DNA Ligation Kit » de Roche [74]. La restauration de la sé-

```

Gène 4H      ACAACAACCTCGGGAATTCGGGAAGGATGGCTACGG· · · · · AAGACATGACCGACATGATGAC
Amorce 1F    -----GAAGGATGGCTACGG· · · · · AAGACATGACCGACATG-----
Amorce 2F    -----
Amorce 3F    -----
Amorce 4F    -----
Amorce corra -----AATTCGGGAAGGATGGCTACGG· · · · ·

Gène 4H      CACCCTGTTTAAAACTATGCAGCTGTTGACCAAGTCGGAACCCACGGCTATGGACGAGGCCACT
Amorce 1F    -----
Amorce 2F    -----
Amorce 3F    -----
Amorce 4F    -----
Amorce corra ······

Gène 4H      AAAACGGCTACTACAATGAAGAATCATCTTCAAAACCTGATGCAGAAGA· · · · · CTAAGAACA
Amorce 1F    -----
Amorce 2F    -----CTTCAAAACCTGATGCAGAAGA· · · · · CTAAGAACA
Amorce 3F    -----GCAGAAGA· · · · · CTAAGAACA
Amorce 4F    -----
Amorce corra ······

Gène 4H      AAG· · · · · AAGACATGACGGACATGGCCACTACGTGGTTCAAAACGATGCAGTTGTTAACGAA
Amorce 1F    -----
Amorce 2F    AAG· · · · · AAGACATGAC-----
Amorce 3F    AAG· · · · · AAGACATGACGGACATGG-----
Amorce 4F    -----
Amorce corra AAG· · · · · ACATGACGGACATGGCCACT-----

Gène 4H      GACCGAGCCCTCGGCCATGGACGAGGCCACGAAGACGGCTACAACCATGAAAAATCATCTGCAG
Amorce 1F    -----
Amorce 2F    -----
Amorce 3F    -----
Amorce 4F    -----GCAG
Amorce corra -----

Gène 4H      AACTTGATGCAAAAA· · · · · GGCGTAGCTCACCACCACCACCACCCTAAAGCTTGGCACTGG
Amorce 1F    -----
Amorce 2F    -----
Amorce 3F    -----
Amorce 4F    AACTTGATGCAAAAA· · · · · GGCGTAGCTCACCAC-----
Amorce corra -----

Gène 4H      CCGTCGTTTTACAACGT
Amorce 1F    -----
Amorce 2F    -----
Amorce 3F    -----
Amorce 4F    -----
Amorce corra -----

```

**FIGURE 3.1** Alignement des diverses amorces mutagéniques sur la séquence de 4H. Ce qui est en gras ne fait pas partie du code pour la protéine. Les bases en blanc surligné représentent les sites AatII.

quence correcte du gène 2HB au niveau du site AatII résultant a été effectuée par mutagenèse dirigée, et ce, grâce aux amorces mutagéniques présentées en figure 3.2.





### 3.2 Expression et purification

Les étapes d'expression et de purification sont délimitées au tableau 3.4. Des bactéries de type *E. coli* DH5 $\alpha$  contenant les divers vecteurs des mutants de MB-1 (tableau 1.1) ont été utilisées. Une préculture (15 mL de milieu Lauri-Bertani (LB) Miller, 15  $\mu$ L ampicilline 100 mg/mL) fut démarrée à 37 °C avec agitation à 250 rpm pour toute une nuit. Après avoir atteint le plateau de croissance (saturation), cette préculture fut ajoutée à un litre de milieu LB Miller contenant de l'ampicilline et supplémenté de glucose (10 %) pour empêcher la production d'amylase, protéine qui interférerait avec la purification subséquente [67]. Ce mélange fut incubé à 37 °C, 250 rpm jusqu'à l'atteinte d'une densité optique (DO) à 600 nm entre 0,3 et 0,5. La transcription fut induite par l'ajout d'IPTG (1 mM) suivie de l'incubation de la culture pour trois heures. Les cellules furent ensuite récoltées par centrifugation à 4000 rpm pendant 40 min.

L'ensemble des étapes de purification est tiré essentiellement de la référence 55, sauf pour quelques modifications. Les cellules précipitées furent d'abord resuspendues à l'aide de tampon de colonne d'amylose (20 mM Tris, 200 mM chlorure de sodium (NaCl), 1 mM d'acide éthylènediaminetétracétique dihydrate (EDTA), 1 mM azoture de sodium (NaN<sub>3</sub>), pH 7,4) froid supplémenté d'un mélange d'inhibiteurs de protéases composé de fluorure d'alpha-phényl méthyl sulfonyl (PMSF), d'EDTA et d'acide (éthylène bis (oxyéthylènenitrolo) ) tétraacétique (EGTA), d'hydrochlorure de benzamidine et de benzamide aux concentrations finales de 0,1 ; 10 ; 10 ; 2 et 2 mM, respectivement. L'ajout de 100 mg de leupeptine et une tablette « Complete Protease Inhibitor Cocktail » de Roche [75] viennent compléter le mélange. Par l'ajout de ces inhibiteurs, on espère minimiser la dégradation des protéines par des protéases relâchées lors de la désintégration subséquente des cellules. Les cellules furent ensuite lysées par 10 coups de 30 s de sonication séparés de pauses de refroidissement de 30 s en utilisant un sonicateur Branson Sonifier 250 réglé à 60 % de puissance. Le mélange fut centrifugé à 12000 rpm, 4 °C pour 20 min

**TABLEAU 3.4**

Détails des étapes de purification.  
Les lignes grasses délimitent les différents jours de manipulations.

Étape	Sous-étape	Conditions expérimentales
Préculture		15 mL LB + 15 µL ampicilline 100 mg/mL + <i>E. coli</i> , 37 °C, 250 rpm, 16 hrs
Culture		1 L LB + 10% glucose + 1000 µL ampicilline 100 mg/mL + 15 mL préculture, 37 °C, 250 rpm jusqu'à $DO_{600nm} = 0,3 - 0,5$
Induction		(Culture à $DO_{600nm} = 0,3 - 0,5$ ) + 1 mL ITPG 1mM, 37 °C, 250 rpm, 3 hrs
Récolte des cellules	Centrifugation	4000 rpm, 40 min, 4 °C
	Resuspension	30 mL tampon colonne amylose + mélange d'inhibiteurs de protéases (voir texte pour détails), 4 °C
Sonication		Puissance 60 %, 10 min (séquence de 30 s sonication, 30 s refroidissement), sur glace / méthanol
Nettoyage de l'échantillon	Centrifugation	12000 rpm, 20 min, 4 °C
	Filtration	Pores de la membrane : 0,22 µm
Chromatographie d'affinité (amylose)	Chargement / lavage	30 mL protéines suivies de 100 mL de tampon de colonne amylose, 4 °C
	Élution	65 mL (tampon de colonne + 10 mM maltose), 4 °C
	Purge	100 mL SDS 0,1 %
Dialyses	Clivage	Pic d'élution amylose + (1 µg/µL facteur Xa / mL pic) contre 2 L tampon clivage (voir texte pour détails), 35 – 40 hrs, 4 °C
	TE	Contre 1 L 10 mM Tris, 1 mM EDTA, 1 mM NaN <sub>3</sub> , pH 8,5 (tampon TE), 2 hrs, 4 °C
Chromatographie d'échange anionique (UNO-Q)	Chargement / lavage	Dialysat suivi de tampon TE, gradient pH 8,4 – 7,7 par étapes de 0,1 unité de pH, 4 °C
	Purge	50 mL tampon TE + NaCl 1 M, 4 °C

afin de déposer les débris cellulaires.

Le surnageant de centrifugation fut filtré pour éliminer les débris cellulaires résiduels et permettre le passage sur chromatographie « fast protein liquid chromatography » (FPLC) ou chromatographie liquide et rapide de protéine (BioLogic DuoFlow System, BioRad). La purification rapide de nos mutants est facilitée par le fait qu'ils sont exprimés en fusion avec une protéine « marqueur » qui est la MBP. Ceci permet d'utiliser une première étape de chromatographie très puissante de type « affinité ». L'échantillon est passé sur une colonne (20mL) de résine présentant des groupements amylose. La protéine MBP, exprimée en fusion avec MB-1, se lie à ces groupements, pendant que toutes les autres protéines sont lavées par passage du tampon. Puis, la fusion « MBP / MB-1 mutant » fut éluée grâce à l'ajout de maltose (10 mM) au tampon de colonne.

La séparation des mutants de MB-1 de la protéine marqueur se fait grâce à la protéase « facteur Xa » qui reconnaît la séquence « NSGR » positionnée entre les deux protéines. Pour ce faire, 1 µg/µL de la protéase fut ajouté par millilitre de protéine éluée. Ce mélange fut ensuite dialysé dans un tube Spectra/Por (limite moléculaire de poids de 3500 Da) contre un tampon permissif à la réaction de clivage (20 mM Tris, 100 mM NaCl, 3 mM chlorure de calcium (CaCl<sub>2</sub>), 0,005 % de dodécyl sulfate de sodium (SDS)) pendant 35 à 40 heures à 4 °C.

Suite au clivage, une deuxième dialyse de l'échantillon se fait contre du tampon 10 mM Tris, 1 mM EDTA (tampon TE) à pH 8,4, 4 °C pour 2 heures. La chromatographie finale est une échangeuse d'anion UNO-Q. Les pI théoriques des différents mutants de MB-1 s'étalent entre 5,17 et 8,49 (tableau 2.3). Il faudra donc utiliser un gradient de pH pour le lavage par souci de répétition exacte des étapes. Dans chaque cas, on devrait obtenir élution des mutants de MB-1 alors que MBP reste fixée sur la colonne.

Après à la chromatographie échangeuse d'ions, les fractions contenant les mutants de MB-1 furent identifiées par électrophorèse sur gel de polyacrylamide contenant du dodécyl sulfate de sodium (SDS-PAGE). Des gels de polyacrylamide-tricine (16 %) furent utilisés pour la séparation puis furent colorés à l'argent. Les fractions contenant les mutants purifiés furent mises en commun puis concentrées par ultra-filtration (Amicon Ultra-15 Centrifugal Filter Devices (cutoff 5000Da) de Millipore) [76]. Les mutants concentrés furent finalement dialysés toute une nuit à 4 °C contre deux tampons de caractérisation, soit un tampon à 128 mM phosphate (pH 6,8) et un tampon borate-phosphate, les deux composants étant à 65 mM (pH 6,8).

### **3.3 Quantification des protéines**

La concentration fut ensuite déterminée par essai à l'acide bicinconinique (test « BCA ») [77]. La protéine utilisée pour calibrer le test est l'albumine de sérum bovin.

### **3.4 Investigation conformationnelle par fluorescence intrinsèque**

La concentration de mutant fut ajustée à 0,1 mg/mL (protéines totales) avec du tampon phosphate. L'échantillon fut équilibré à la température de la pièce pendant une heure avant la prise de mesure. La protéine fut excitée à 280 nm dans un fluorimètre Cary Eclipse de Varian et le spectre d'émission fut enregistré entre 290 et 400 nm. Le spectre obtenu fut corrigé pour l'émission du tampon.

### **3.5 Investigation conformationnelle par dichroïsme circulaire**

La concentration de mutant fut ajustée à 0,4 mg/mL (protéines totales) avec du tampon phosphate. L'échantillon fut équilibré à 20 °C pendant 20 min avant la prise de mesure avec un spectropolarimètre Jasco J-720. Les détails des conditions sont présentés au tableau 3.5. Le signal DC obtenu fut converti

**TABLEAU 3.5**  
Détails des lectures de dichroïsme circulaire

Paramètres	Valeurs
Début des lectures	260 nm
Fin des lectures	180 nm
Résolution	1 nm
Vitesse de lecture	20 nm/min
Nombre d'accumulations	10
Réponse	8 s
Largeur de bande	1,0 nm
Sensibilité	20 mdeg

en  $\Delta\epsilon_{MRW}$ , ou MolCD selon le programme Jasco, mesure qui représente la différence de coefficient d'extinction molaire par mole de résidus. Le pourcentage de structures secondaires fut calculé via CDSSTR en utilisant les paramètres par défaut [ce programme est disponible à partir de la référence 78].

### 3.6 Dénaturation thermique

L'impact des mutations sur la stabilité a été évalué par DC. On a mesuré la thermostabilité du mutant lors d'un chauffage progressif de 10 à 90 °C. Ce chauffage fut obtenu grâce à un bain thermostaté Neslab RTE-111 contrôlé par le programme du spectropolarimètre Jasco permettant une augmentation de 30 C par heure. La température indiquée par l'appareil Jasco fut corrigée à l'aide de mesures de thermocouple prises directement à l'échantillon. Il fut assumé que la perte de signal DC à 222 nm reflétait les changements conformationnels induits par la chaleur. Afin de quantifier une différence de

thermostabilité entre les protéines, la température de fusion fut calculée tel que suggéré par Pace [79]. Pour ce faire, il a été assumé que la dénaturation était réversible, coopérative et unimoléculaire.

### 3.7 Dégradation protéolytique

Ces manipulations se sont déroulées selon Krishnamoorthy [80] en incluant les modifications données dans Morrison [81]. Ce protocole fait appel au mélange Pronase E, qui est un assortiment de protéases (E.C. 3.4.24.31) de *Streptomyces griseus*. La concentration des mutants est d'abord ajustée à 0,2 mg/mL (protéines totales) avec du tampon borate-phosphate dans un volume final de 100  $\mu$ L puis l'échantillon est stabilisé à 39 °C. Deux échantillons sont préparés par mutant, dont un qui sera le contrôle sans protéases. On ajoute ensuite du mélange Pronase E à l'échantillon « avec protéases » pour obtenir une concentration finale de 0,211 mg/L. Une aliquote initiale, au temps 0 min, est d'abord prise, puis l'incubation à 39 °C continue. Des aliquotes sont ainsi prises aux temps 5, 10, 20, 30, 45, et 60 min. Le tout est fait en triplicata.

Les réactions furent stoppées par ajout de tampon 2 % SDS et chauffage à 95 °C pour 5 min. L'analyse des dégradations s'est faite par SDS-PAGE suivi d'une coloration au nitrate d'argent. La quantité de protéines résiduelles après digestion fut déterminée en mesurant la densité optique des bandes (appareil FluorChem et programme d'analyse d'image AlphaEaseFC d'Alpha Innotech). Afin de pouvoir corriger pour des différences expérimentales lors des différentes étapes de cette méthode, le cytochrome c de cœur de cheval a été utilisé comme protéine « contrôle » (dégradation et coloration).

## CHAPITRE 4

### RÉSULTATS

#### 4.1 Constructions génétiques

Au début de ce projet, il n'y avait que les constructions génétiques de MB-1TrpHis, 4H-EV-His et 6H-EV-His puis les produits PCR de 2HA-His et 2HB-His de disponible. Une grande série de manipulations a dû être effectuée sur ces ADN initiaux pour obtenir les sept constructions nécessaires à cette étude. Ces étapes furent délimitées dans la section « Stratégie » de cet ouvrage, puis furent résumées au tableau 2.2.

Premièrement, les produits PCR de 2HA-His et 2HB-His devaient être insérés dans le même vecteur d'expression que les autres mutants, soit pMAL-c2. Ces clonages se sont avérés inefficaces, même après plusieurs modifications au protocole (ratio vecteur/insert, temps et température d'incubation) qui devaient augmenter les chances de réussite. La nouvelle stratégie se composant de rogner le vecteur contenant le gène de MB-1TrpHis fut donc envisagée puis effectuée, résultant en l'obtention des constructions voulues. Ce processus comporte de trois (2HA-His et 2HA-EV-His) à quatre (2HB-His) étapes au lieu d'une seule pour le clonage direct.

Deuxièmement, le variant « EV » de 6H-His a dû être créé. Ceci n'implique qu'une étape, mais le besoin d'obtenir une construction ayant les deux sites (N<sub>44</sub>L<sub>45</sub> et N<sub>93</sub>L<sub>94</sub>) modifiés simultanément compliqua la recherche de clones positifs.

Finalement, la similitude de pl entre la MBP et les différents variants à queue d'histidine étudiés au début de ce projet a occasionné des difficultés de purification (voir la section « Stratégie »). Il fut donc décidé de retourner au système de purification initial sans queue d'histidine pour tous les variants. Ceci



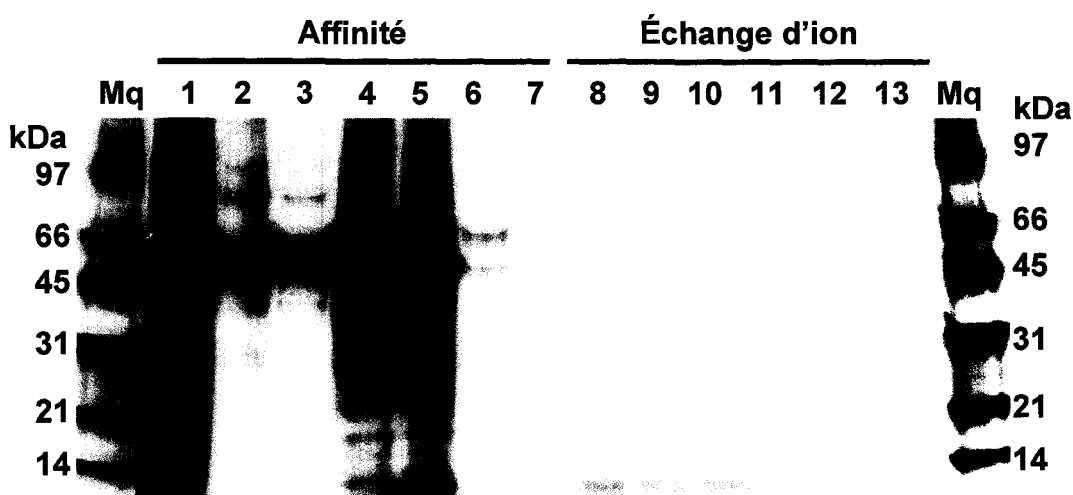
implique une étape supplémentaire de mutation pour toutes les constructions génétiques utilisées lors de ce projet. En tout et partout, il a fallu entre une et cinq étapes de modifications pour obtenir chacune des constructions finales.

## **4.2 Expression et purification**

La purification des protéines nécessaires à ces travaux fut suivie par analyse SDS-PAGE. Les analyses de 4H-EV et 2HA sont deux exemples typiques des différentes purifications effectuées lors de ce projet.

Les différents essais de purification de 4H-EV se sont bien déroulés, avec suffisamment de protéines pures recueillies à la fin pour déterminer les caractéristiques physiques de ce mutant. La figure 4.1 présente plusieurs fractions récoltées au cours de la cette purification, qui se fait en deux étapes majeures; la chromatographie d'affinité (partie gauche de la figure) et la chromatographie d'échange d'ions (partie de droite). La première chromatographie utilise l'affinité pour l'amylose que possède la protéine MBP greffée au mutant de 4H. Une grande quantité de cette protéine fusion (environ 60 kDa) est concentrée à l'élution (puit 4). S'en suit le clivage (puit 5) de cette dernière pour produire MBP (45 kDa) et, dans le cas présent, 4H-EV (environ 14 kDa). La différence entre les *pI* (tableau 2.3) de la MBP et de 4H-EV permet de séparer les protéines par échange d'ions. Cette séparation s'avère très efficace, tel qu'en témoignent les puits 8 à 13, où l'on perçoit 4H-EV et aucune trace de MBP ou de protéine fusion résiduelle.

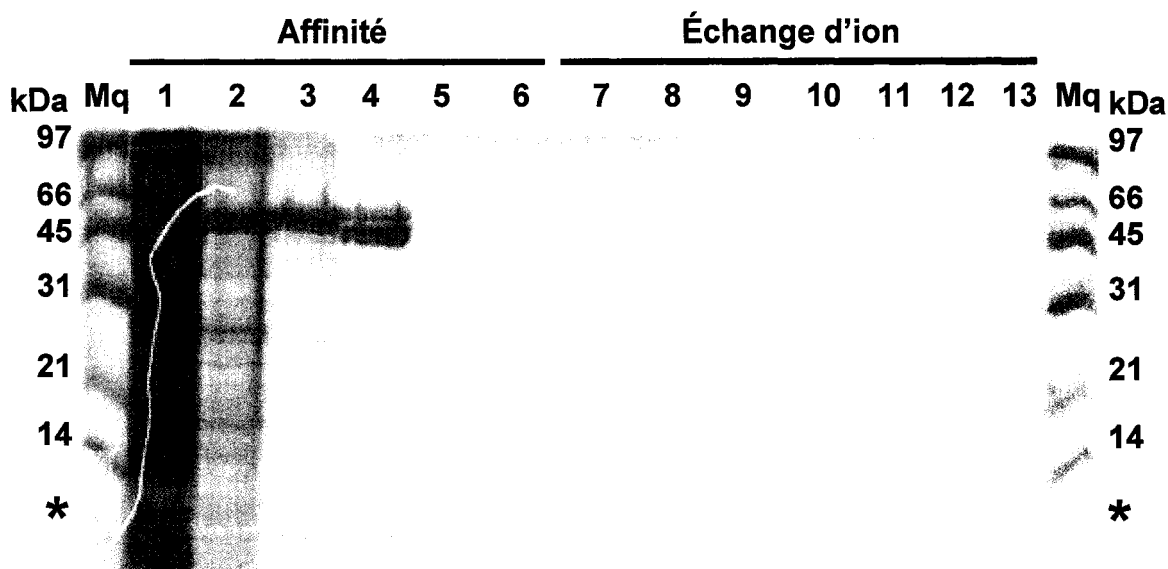
En ce qui concerne la purification de 2HA (figure 4.2), les résultats sont tout autres. La chromatographie d'affinité fonctionne bien, avec accumulation de protéine fusion (environ 56 kDa) après l'élution (puit 3). De plus, il semble que le clivage de la protéine fusion (puit 4) ait partiellement fonctionné, avec libération de MBP et possiblement 2HA (vers 11 kDa). Cependant, aucune quantité de 2HA n'est présente lors de la dernière étape de purification, soit la



**FIGURE 4.1** SDS-PAGE coloré au nitrate d'argent de différentes fractions recueillies au cours de la purification de 4H-EV. Les marqueurs de poids moléculaire (Mq), en bouts de gel, ont des tailles de 97, 66, 45, 31, 21 et 14 kDa respectivement, de haut en bas. Les échantillons de chromatographie d'affinité sont aux puits 1-7, alors que les échantillons d'échange d'ions sont aux puits 8-13. Puits; 1 = premier volume mort, 2 = deuxième volume mort, 3 = troisième volume mort, 4 = élution avant clivage, 5 = élution après clivage, 6 = première purge, 7 = deuxième purge, 8 – 13 = fractions de l'échange d'ions; étape où l'élution est attendue.

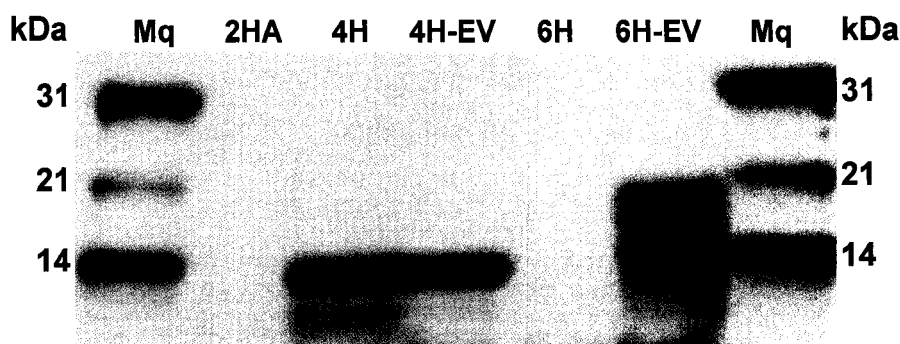
chromatographie d'échange d'ions (puits 7 à 13), et ce, malgré plusieurs différents essais de purification et d'ajouts d'inhibiteurs de protéases.

Quelques légères difficultés sont survenues lors des purifications. Pour chaque mutant, on note l'impossibilité de capter toute la protéine fusion, qui passe en partie dans le volume mort de la chromatographie d'affinité (figures 4.1 et 4.2 puits 1 à 3), et ce, même si le volume de résine est permissif à la capture totale. On remarque aussi le clivage incomplet de la protéine fusion pour chaque mutant (figure 4.2, puits 4 et 5). Ce phénomène contribue à diminuer la quantité de mutant recueilli à la fin de la purification. De plus, il y a constamment présence de fragments de dégradation (8-10 kDa) avant et après le clivage de la protéine fusion.



**FIGURE 4.2** SDS-PAGE coloré au nitrate d'argent de différentes fractions recueillies au cours de la purification de 2HA. Les marqueurs de poids moléculaire (Mq), en bouts de gel, ont des tailles de 97, 66, 45, 31, 21 et 14 kDa respectivement, de haut en bas. Il y a un fragment de dégradation présent en fin de marqueur, indiqué par une astérisque. Les échantillons de chromatographie d'affinité sont aux puits 1-6, alors que les échantillons d'échange d'ions sont aux puits 7-13. Puits; 1 = premier volume mort, 2 = deuxième volume mort, 3 = élution avant clivage, dilué 1/6, 4 = élution après clivage, dilué 1/6, 5 = première purge, 6 = deuxième purge, 7 – 13 = fractions de l'échange d'ions; étape où l'élution est attendue.

La figure 4.3 montre les résultats finaux de certaines purifications. Il s'agit d'une électrophorèse des différents mutants selon la quantification BCA. Les tailles attendues sont d'environ 11, 14 et 17 kDa pour les divers mutants de 2H, 4H et 6H respectivement. Les bandes de protéines intactes de 4H, 4H-EV et 6H-EV sont près des tailles attendues. D'ailleurs, la protéine 4H-EV présente un profil intéressant de pureté. Cependant, on remarque une fragmentation chez 4H et 6H-EV. Pour 4H, le fragment présenté ici est identique en taille à celui détecté lors de projets antérieurs [54]. En ce qui concerne 6H-EV, on note la présence de plusieurs fragments de tailles différentes s'étalant entre approximativement 16 et 14 kDa. On note aussi l'impossibilité d'accumuler 2HA



**FIGURE 4.3** SDS-PAGE des différentes purifications des mutants, gel coloré au nitrate d'argent. Les marqueurs de poids moléculaire (Mq), en bouts de gel, représentent 31, 21 et 14 kDa respectivement, de haut en bas.

(2HA-EV et 2HB aussi, résultats omis par souci de simplicité) et 6H. Aucune bande n'est présente aux tailles attendues pour ces variants. Ce manque d'accumulation des mutants à deux hélices et de 6H ainsi que les divers fragments retrouvés chez 4H et 6H-EV peuvent indiquer un manque de résistance à la dégradation protéolytique. Cette plus faible résistance découle possiblement d'une lacune dans la stabilité conformationnelle des variants.

Connaissant toutes ces informations, la caractérisation des protéines non accumulées (toutes les 2H et la 6H) est impossible. Cela n'est pas le cas pour 4H et 4H-EV, qui sont suffisamment purs pour les analyses subséquentes. De plus, une caractérisation de 6H-EV sera effectuée, mais elle sera partielle étant donné le faible degré de pureté de cette protéine.

### 4.3 Investigation conformationnelle par fluorescence intrinsèque

La vérification de l'émission fluorescente des chaînes latérales de certains acides aminés (aromatiques) peut fournir un indice quant à la structure tertiaire d'une protéine. Un résidu fluorophore judicieusement positionné dans la molécule permet de discerner le degré de compaction de la région avoisinant ce

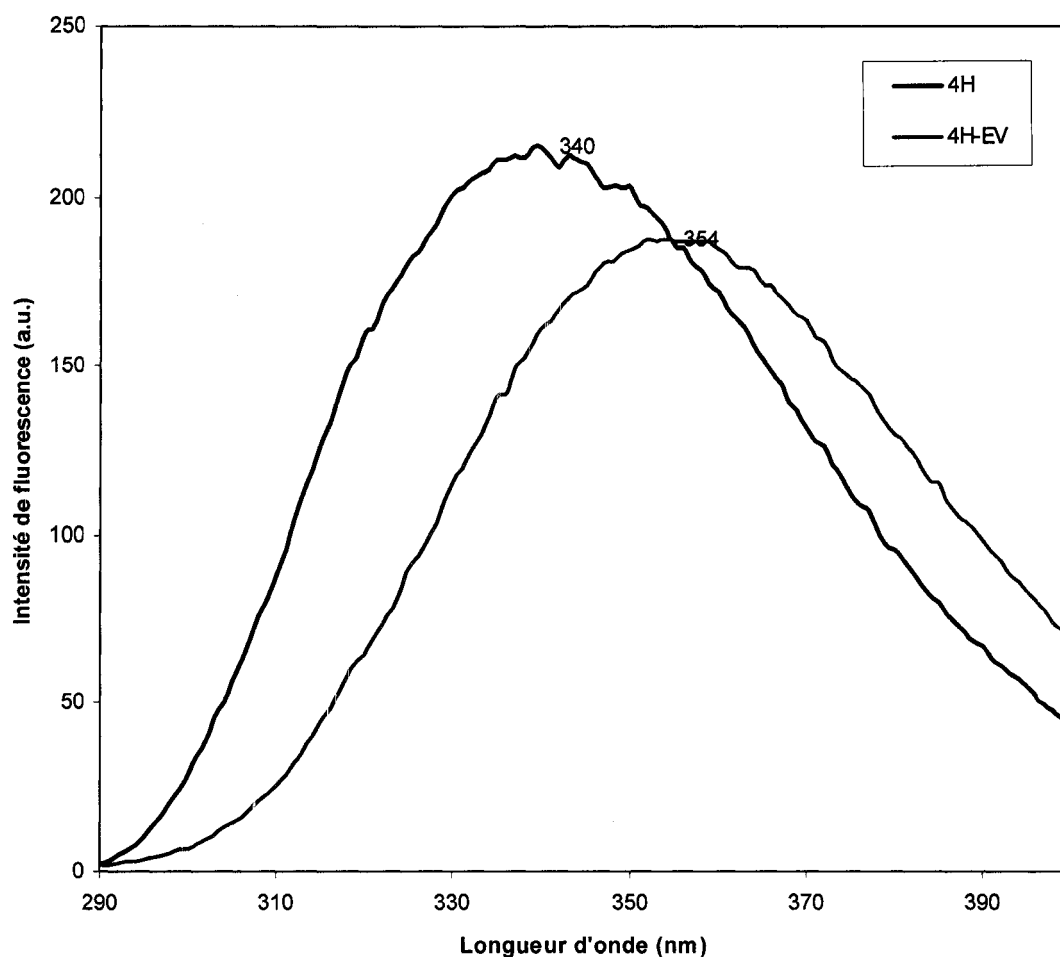
résidu. Chez 4H et ses divers mutants, le fluorophore W devrait être enfoui au cœur de la protéine. Vu la petite taille de ces molécules, il est possible de s'avancer sur la compaction de ces structures complètes selon les résultats de fluorescence du W.

L'émission de fluorescence du W lorsqu'excité à son maximum d'absorption (280 nm) est caractéristique selon le degré d'exposition de sa chaîne latérale. Un W bien enfoui au cœur d'une structure protéique émettra de la fluorescence à un niveau maximal autour de 325 nm, alors qu'un W totalement exposé à la surface d'une protéine émettra avec un maximum à 355 nm [83]. Connaissant ces valeurs bornes, il est possible de se prononcer sur le degré d'exposition du W chez 4H et 4H-EV et d'en déduire une idée de la compaction de la structure entière. Le maximum d'émission de fluorescence chez 4H se produit à 340nm (figure 4.4). Cette valeur indique un W plus ou moins exposé. Cependant, après modification du site N<sub>44</sub>L<sub>45</sub> vers E<sub>44</sub>V<sub>45</sub> pour obtenir 4H-EV, le maximum d'émission fluorescente est déplacé vers 354 nm, ce qui indique une exposition totale du W. Ceci est un indice semblant dévoiler un relâchement de la structure tertiaire chez 4H-EV.

#### **4.4 Investigation conformationnelle par dichroïsme circulaire**

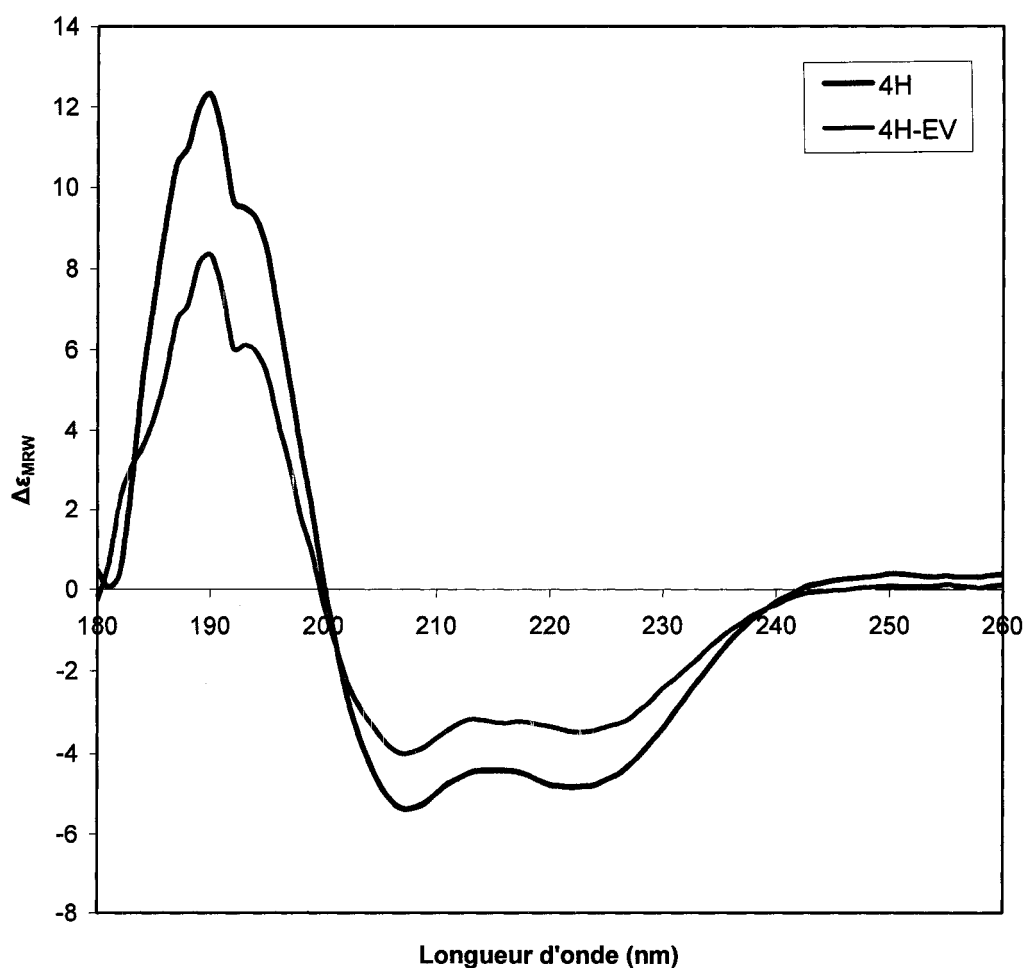
Les essais de dichroïsme circulaire sur une protéine permettent de spécifier la répartition des différentes structures secondaires au sein de la molécule. Lorsque plusieurs mutants de la même protéine sont à l'étude, tel que dans ce projet, il est possible de discerner des changements conformationnels entre les différents variants.

Les spectres DC en figure 4.5 sont typiques d'une protéine comportant une grande quantité d'hélices  $\alpha$ . Ils permettent de calculer les pourcentages de structures secondaires des protéines, qui sont présentés au tableau 4.1. On no-



**FIGURE 4.4** Spectre d'émission fluorescente de 4H et 4H-EV lorsqu'excités à 280 nm.

te une disparition (14,7 %) de structures hélicales chez 4H-EV lorsqu'on compare ce variant avec la protéine 4H. Cette diminution d'hélices semble être transférée quasi équitablement aux feuillets  $\beta$ , aux connecteurs et aux structures désordonnées, représentant des augmentations de 4,5, 4 et 5,7 % respectivement. Il semblerait donc que la modification à la structure tertiaire de 4H-EV passe par un réarrangement (feuillets  $\beta$ ) et une déstructuration (connecteurs et aux structures désordonnées) des structures secondaires de 4H.



**FIGURE 4.5** Spectre de dichroïsme circulaire de 4H et 4H-EV à 20 °C.

**TABLEAU 4.1**

Pourcentages de structures secondaires de 4H et 4H-EV calculés via CDSSTR.

Nom du mutant	Pourcentage (%) de			
	Hélices $\alpha$	Feuillets $\beta$	Coudes / connecteurs	Structures désordonnées
4H	57,4	13,9	11,4	17,5
4H-EV	42,7	18,4	15,4	23,2
4H-EV vs 4H	-14,7	4,5	4	5,7

## 4.5 Dénaturation thermique

Les essais de DC sont aussi utiles pour étudier la dénaturation des protéines. Pour suivre la dénaturation thermique de 4H et 4H-EV, la perte de signal DC à 222 nm (représentatif de structures hélicales) fut enregistrée à la figure 4.6. On note que les courbes sont presque linéaires, ce qui témoigne d'une dénaturation peu coopérative. Cela signifie que les protéines sont plutôt sous forme de globules fondus que sous forme bien compactes. De plus, on note une disparition plus rapide du signal lorsqu'on chauffe 4H comparé à 4H-EV.

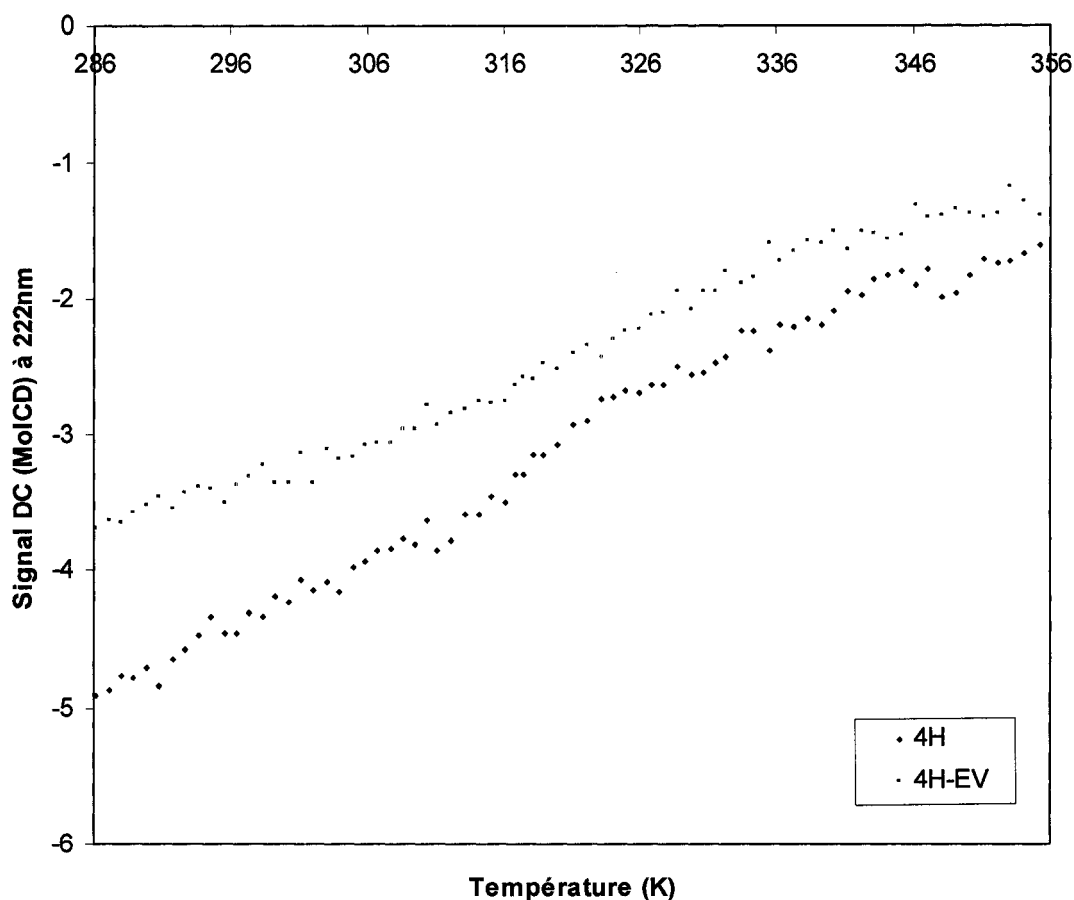
Il est tout de même possible de discerner des sections de pré- et post-transitions pour chaque courbe, ce qui permet de calculer le changement d'énergie libre ( $\Delta G$ ) de dénaturation. Selon Pace *et al* [79], le  $T_m$  d'une protéine est l'intercepte à l'axe des abscisses du graphique de  $\Delta G$  de dénaturation en fonction de la température (figure 4.7). Le  $T_m$  de 4H est de 44,43 °C alors que celui de 4H-EV est de 53,73 °C, ce qui représente une augmentation d'un peu plus que 9 °C pour cette dernière. La dénaturation thermique de 4H serait donc plus facile à effectuer.

## 4.6 Dégradations protéolytiques

Le but ultime de ce projet est de trouver un mutant de MB-1 suffisamment stable pour passer le rumen des vaches laitières puis être digéré au bon emplacement dans l'animal. Les essais de dégradations protéolytiques permettent de vérifier la résistance aux protéases des divers mutants. Dans le cas présent, le mélange Pronase E fut utilisé, car il mime les protéases ruminales.

Il est possible d'analyser la résistance au Pronase E pour 4H, 4H-EV ainsi que 6H-EV dans ce projet. Il est vrai que 6H-EV présente des fragments de dégradation avant même l'analyse, mais la vérification de la perte de protéines

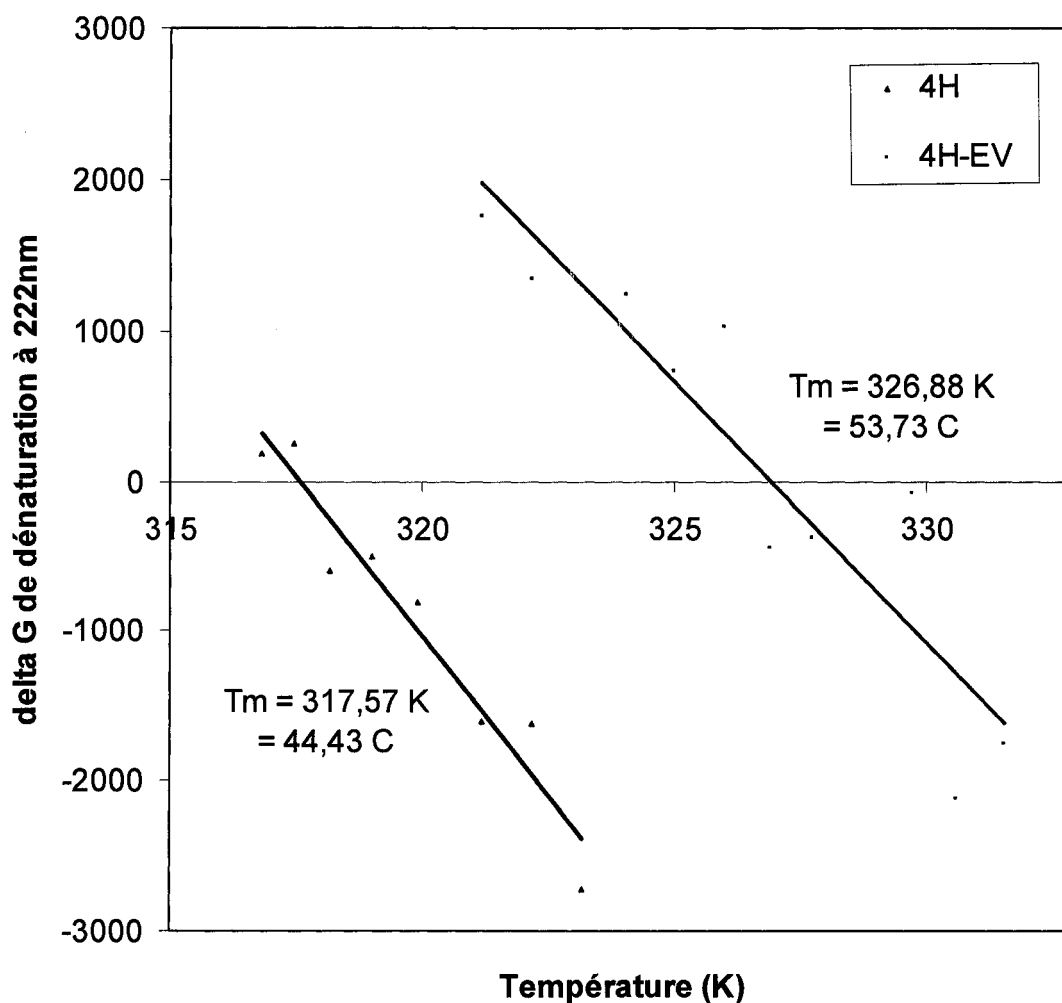




**FIGURE 4.6** Suivit de la dénaturation thermique de 4H et 4H-EV par dichroïsme circulaire à 222 nm.

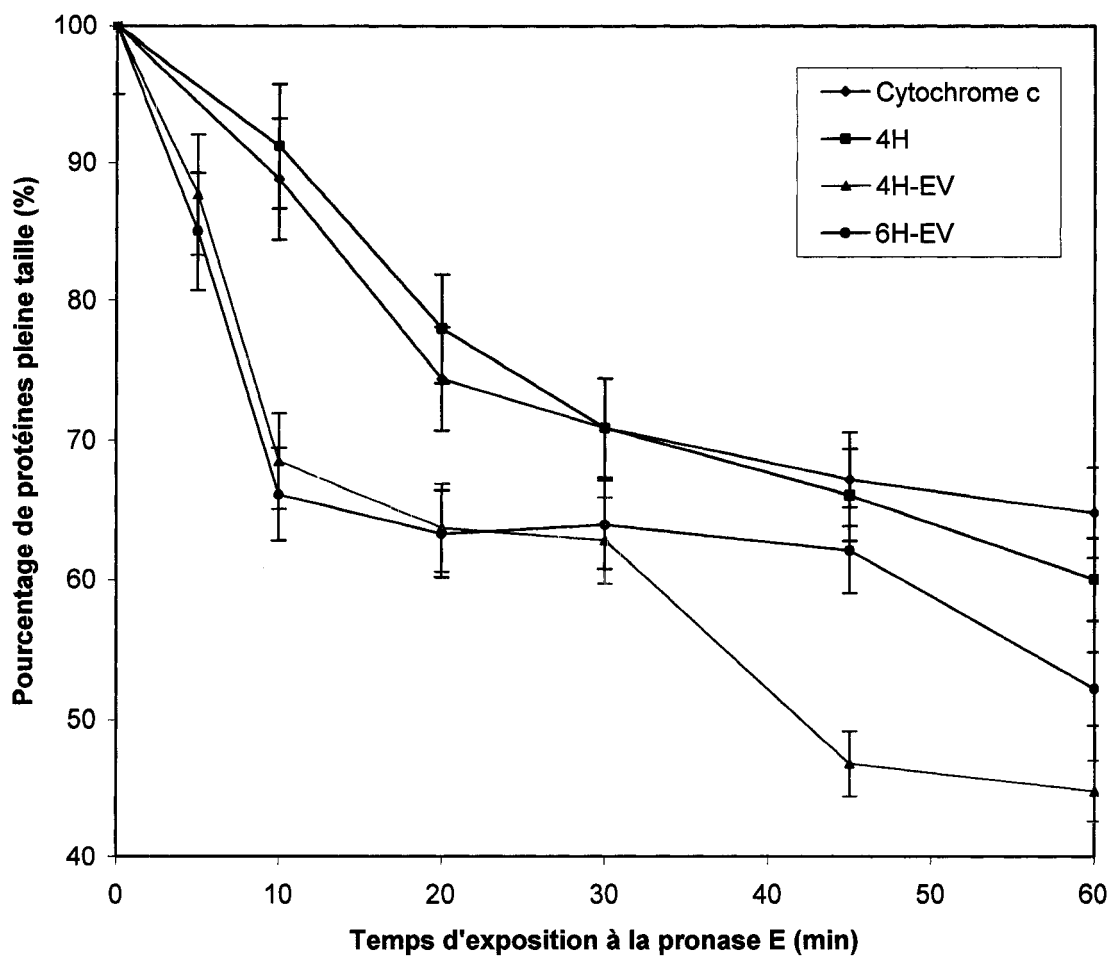
pleines tailles est néanmoins possible. La figure 4.8 présente les courbes de dégradations protéolytiques des mutants mentionnés ci-haut. On remarque que la dégradation de 4H est similaire à celle du cytochrome c, standard de cette expérience, et que celles de 4H-EV et de 6H-EV sont semblables. Cependant, et ce contre les attentes, les deux variants présentant la modification NL vers EV semblent moins résistants au Pronase E.

La quantité de protéine pleine taille résiduelle de chaque mutant après 10 minutes d'exposition au Pronase E est diagnostique. Le variant 4H suit la tendance du standard cytochrome c, avec approximativement 90 % de protéines



**FIGURE 4.7**  $\Delta G$  de dénaturation à 222 nm pour 4H et 4H-EV.

intactes, tel que dans les expériences de Gagnon *et al* [55]. Les résultats sont bien différents pour 4H-EV et 6H-EV, qui ne conservent qu'environ 70 % des protéines pleine longueur. Il semble que le site « EV » soit possiblement attaqué par ce mélange particulier de protéases. De façon surprenante, cette différence de dégradation semble s'atténuer progressivement entre les temps de 10 minutes et 30 minutes. Cependant, la dégradation de 4H-EV s'intensifie de manière importante à partir de ce point, avec une quantité de protéines pleines tailles d'environ 45 % alors que toutes les autres protéines conservent toujours approximativement 60 – 70 % de leurs molécules intactes.



**FIGURE 4.8** Dégradations protéolytiques (Pronase E) à 39°C des différents mutants de MB-1 à 0,2 mg/mL.

Au bout d'une heure d'exposition, la tendance de résistance au Pronase E semble bien établie; cytochrome c, 4H, 6H-EV et 4H-EV, du plus au moins résistant respectivement.

## **CHAPITRE 5**

### **DISCUSSION ET CONCLUSIONS**

#### **5.1 Constructions génétiques pour à la production des mutants 2H**

Tel qu'il fut mentionné dans la section « Stratégie », les différents essais de clonage des gènes des mutants 2H dans le vecteur pMAL-c2 ont été infructueux. Il fut supposé que la cause de ce problème était le manque de collisions entre vecteurs et inserts digérés. D'autre part, il fut aussi considéré que ces collisions avaient bel et bien lieu, mais que l'intégrité des bouts cohésifs du vecteur ou de l'insert soit atteinte par la présence possible de DNases contaminantes.

Le processus de rognage du vecteur comportant le gène codant pour quatre hélices pour n'en faire que deux fut choisi en lieu du clonage initial. Les raisons sont les suivantes. Les chances de collisions des bouts cohésifs du vecteur sur lui-même sont meilleures (proximité). De plus, les chances de contaminations sont amoindries par élimination d'étapes de manipulations. Ce changement a effectivement permis d'obtenir les constructions voulues.

#### **5.2 Expression et purification**

##### **5.2.1 Obstacles rencontrés**

La production de la protéine fusion initiale s'est toujours bien déroulée. La comparaison des SDS-PAGE contenant des fractions induites et non induites des souches transformées permettait de bien discerner la présence de fusion à l'induction. Cependant, c'est au niveau des purifications que divers problèmes sont apparus.

Le premier obstacle rencontré lors des divers essais de purification était l'adhésion inefficace de la protéine fusion lors de la chromatographie d'affinité, causant une perte de rendement initial de purification (figures 4.1 et 4.2). Il y a trois raisons possibles pour expliquer cette perte directe de protéine fusion. Premièrement, il est possible que la souche utilisée ait produit de l'amylase dans le milieu de culture choisi. L'amylase peut affecter la purification en empêchant la protéine fusion d'adhérer à la colonne d'affinité en coupant l'amylose greffée à la résine ou en produisant du maltose qui vient éluer la protéine fusion de la colonne [67]. Ce problème aurait dû être réglé avec l'ajout de glucose au milieu de culture, qui réprime l'expression de l'amylase.

Deuxièmement, le nombre de fois que la même résine fut utilisée est crucial. La présence possible d'amylase dans l'extrait cellulaire viendrait détériorer graduellement la qualité d'affinité de la résine. Il est donc conseillé de changer la résine à chaque trois à cinq utilisations. Lors des présentes purifications, la résine fut changée à toutes les deux expériences pour éviter une trop grande perte d'efficacité. Finalement, la sonication des cellules pour en extraire les fusions causerait aussi la libération de composants cellulaires qui interféreraient possiblement avec l'attachement de la fusion sur la résine. La sonication est le moyen de choix pour l'extraction de la protéine fusion selon le manuel du système d'expression [67]. Néanmoins, c'est probablement ce dernier point qui vient influencer la diminution d'efficacité de la chromatographie. Cependant, la quantité de fusion captée fut suffisante pour la suite des étapes de purification.

Un second obstacle encouru lors des purifications était l'étape de clivage de la protéine fusion pour obtenir la séparation de la MBP du mutant de MB-1 y étant accroché. On note, à la figure 4.2, que le clivage n'est pas complet et qu'il reste une quantité appréciable (environ 10 %) de la protéine fusion intacte. Plusieurs modifications ont été tentées pour rendre cette étape plus efficace (ajustements de temps, de température, de quantité de facteur Xa et de

détergents). Ces modifications avaient pour but de légèrement perturber la structure de la protéine fusion au cas où le site de clivage serait inaccessible. Cependant, le clivage ne fut amélioré que très peu avec les différentes modifications seules ou même en combinaisons. Néanmoins, la quantité de mutant de MB-1 relâchée par le clivage inefficace était suffisante pour la poursuite de la purification.

Un autre problème rencontré fut la présence de fragments de dégradation de la protéine fusion. Ces fragments, présents peu importe le mutant exprimé, font entre 8 et 10 kDa (figure 4.2), et peuvent être le résultat de l'expression elle-même et de la sonication. La fusion est exprimée dans le cytoplasme de son hôte, puis est exposée aux protéases de ce milieu, subissant une dégradation primaire. Lors de la sonication, les protéases du périplasme et de la membrane externe viennent s'ajouter au mélange, complétant la dégradation initiale de la protéine fusion [68]. Cette dégradation est faible, mais joue quand même sur la quantité totale de protéine amassée. Plusieurs différents essais d'ajouts d'inhibiteurs de protéases n'ont pas permis d'amoindrir ce phénomène.

Le plus important problème rencontré est sans doute les purifications elles-mêmes des protéines 2HA, 2HA-EV, 2HB et 6H. Il fut catégoriquement impossible de récupérer ces mutants à la fin de leurs purifications, et ce, malgré plusieurs différentes modifications au protocole. La purification des mutants 2H par échange ionique fut prévue comme étant difficile, car les pI de ces protéines se rapprochent beaucoup de celui de la MBP. Contrairement, la purification par échange ionique de 6H devait être plutôt facile, vu la grande différence de pI entre cette protéine et la MBP (voir le tableau 2.3). Cependant, aucune trace de ces mutants ne fut retrouvée par SDS-PAGE ni après la chromatographie d'échange d'ions, ni même après l'étape de clivage. Même l'ajout de plus d'inhibiteurs de protéases immédiatement avant la sonication n'a pas pu résoudre ce problème.

### 5.2.2 Différence majeure entre les résultats antérieurs et présents

L'impossibilité d'accumuler les protéines 2H ainsi que la 6H lors de ce projet est surprenante. Des études antérieures [57,58] ont démontré que les protéines à deux ou six hélices sont autant sinon plus stables vis-à-vis les protéases que MB-1TrpHis. Pourquoi n'y a-t-il donc pas de protéine récupérée dans le projet actuel? La réponse réside possiblement dans les conditions d'expression de la protéine lors des différentes caractérisations.

Le projet « alanines de cœur » a permis la découverte des protéines stables à 2H via un criblage d'une banque de variants de MB-1TrpHis par expression phagique. Le criblage mettait à l'épreuve des variants toujours fusionnés au phage, ce qui peut expliquer la stabilisation, on le sait maintenant, faussement attribuée au nombre diminué d'hélices. La stabilisation provient donc peut-être du fait que les protéines 2H se collent au phage d'une manière qui cache ses sites sensibles au mélange Pronase E. Il est possible que le repliement des 2H produites seules soit incomplet, ou carrément inexistant.

L'autre projet, ledit « NL », a permis de supposer que 6H était plus stable que 4H. De la même manière que le précédent projet, le 6H fut aussi caractérisé sous forme de fusion, cette fois-ci à MBP plutôt qu'à une protéine de capsule de phage. Ici encore, il est possible que le repliement soit inadéquat, et que la séquence adhère à la MBP, augmentant ainsi sa stabilité.

Un fait très intéressant vient appuyer la théorie de stabilisation par fusion utilisée ci-haut. Les domaines TRP de protéines naturelles étudiés par Main et ses collaborateurs fonctionnent souvent en tant que médiateur d'interactions protéiques [66]. Il est donc fort possible que les mutants 2H et 6H, similaires en plusieurs points au motif TPR, soient collés à la capsid de phage ou à la MBP, augmentant ainsi leur stabilité globale. Dans ce projet, toutes les protéines ont été étudiées après clivage de la protéine fusion. L'effet stabilisant que la MBP

apportait aux variants de MB-1 n'était donc pas présent puis les 2H et 6H sont devenus plus susceptibles aux attaques des protéases dans le milieu. De plus, il est reconnu [67] que la modification conformationnelle du site de clivage nuit à la coupure de la protéine fusion. L'adhésion des mutants à la MBP pourrait donc aussi expliquer la difficulté encourue avec cette étape de la purification.

### 5.3 Effet bénéfique de la mutation « EV »

Un phénomène intéressant se produit lorsque la mutation « EV » est présente dans la séquence. Il est attendu que cette modification aide à la résistance de MB-1TrpHis dans le milieu bactérien de production. C'est aussi le cas de certaines des protéines présentées lors de ces travaux. La protéine initiale (4H) présente un fragment de dégradation à la fin de la purification, alors que ce fragment est inexistant lorsque N<sub>44</sub>L<sub>45</sub> est changé pour E<sub>44</sub>V<sub>45</sub> (4H-EV).

De plus, il semble que cette modification soit bénéfique pour 6H. Lorsque la mutation « EV » n'est pas présente, l'accumulation de la protéine est impossible. Inversement, quand cette modification est apportée à la séquence, une accumulation de protéine pleine taille est possible. Cependant, cette accumulation est faible, puis elle est accompagnée de plusieurs fragments de dégradation. Il existerait donc d'autres sites de reconnaissance des protéases d'*E. coli* au sein de la séquence, lorsqu'on est sous forme de 6H, qui deviennent suffisamment exposés pour permettre la digestion. Ceci pousse à croire que l'arrangement de 6H n'est pas exactement sous la forme attendue de six hélices, mais plutôt une structure, possiblement différente, plus lâche et souple. La vérification exacte, par fluorescence et DC, de la structure de 6H-EV n'a pu être effectuée compte tenu du manque de pureté de l'échantillon; la protéine pleine taille est contaminée de fragments de dégradation qui ont coélué lors de la purification finale.



Cependant, le même effet stabilisant de « EV » n'a pas été noté chez 2HA, ce qui ajoute à l'hypothèse de mauvais repliement de ce variant posé à la sous-section précédente.

## **5.4 Caractérisations**

### **5.4.1 Investigation conformationnelle par fluorescence intrinsèque**

Les résultats de fluorescence (position du maximum d'émission) suggèrent que 4H-EV est sous une forme plus lâche que celle des 4H. Un second indice appuyant cette théorie peut être discerné grâce à ces essais de fluorescence. On y note, en plus du changement d'emplacement de maximum entre les protéines 4H et 4H-EV, la différence d'intensité du signal, celui de 4H-EV étant légèrement plus faible. Il est possible que le W de 4H-EV subisse une plus grande atténuation (« quenching ») de sa fluorescence. Le phénomène d'atténuation survient lorsque des fluorophores entrent en contact avec des atténuateurs. Lorsqu'un fluorophore entre en collision avec des atténuateurs, il leur transfère son énergie sans émission de fluorescence, ce qui diminue le signal total. Des exemples d'atténuateurs sont l'oxygène dissout, le citrate et certains composants des protéines elles-mêmes, soient les groupements carboxyliques non ionisés, les amines et les groupements guanidines de l'arginine [82]. Dans le cas présent, la plus grande malléabilité de 4H-EV permettrait un plus grand nombre de collisions entre le W et les atténuateurs intrinsèques, diminuant ainsi l'intensité de fluorescence. Le phénomène d'atténuation peut aussi expliquer la plus grande intensité de signal de 4H lorsqu'on considère le(s) fragment(s) de dégradation chez ce variant. Ces fragments peuvent contenir un W (probablement enfoui selon la position du maximum d'émission), et donc émettre de la fluorescence. Les atténuateurs intrinsèques normalement retrouvés au sein de la protéine pleine taille seraient trop distant pour manifester leur influence au niveau de ces fragments. Le signal en est d'autant plus augmenté.

### 5.4.2 Investigation conformationnelle par dichroïsme circulaire

Les résultats de fluorescence ont permis de déterminer un relâchement de la structure tertiaire de 4H-EV lorsque comparé à 4H. Les résultats de DC permettent de déterminer si ce changement structurel affecte aussi l'intégrité des structures secondaires. Ces derniers résultats ont déterminé que la modification « EV » réorganise 4H vers une structure comportant moins d'hélices et plus de feuillets. De plus, ces résultats démontrent une augmentation en coudes, connecteurs et structures désordonnées, appuyant ainsi les essais de fluorescence suggérant un relâchement de la structure tertiaire. Il est impossible de se prononcer sur l'impact de la présence de fragment(s) chez 4H sur le signal, car on ne peut déterminer avec certitude si ce(s) peptide(s) a(ont) une structure apériodique capable d'influencer ce dit signal.

Il est possible que cette réorganisation change la stabilité thermique de 4H-EV lorsque comparée à celle de 4H. Des essais plus poussés de dichroïsme circulaire ont permis de déterminer si ce changement avait lieu, et s'il était positif ou négatif. Ils sont documentés à la section suivante.

### 5.4.3 Dénaturation thermique

Les résultats de dénaturation thermique ont indiqués que 4H-EV serait plus stable que sa consœur ne portant pas la modification N<sub>44</sub>L<sub>45</sub> vers E<sub>44</sub>V<sub>45</sub>. Ce résultat est inattendu si on ne tient compte que de la compaction de la structure. Les essais de fluorescence et de DC indiquent une plus grande flexibilité pour 4H-EV que pour 4H. Les protéines plus lâches devraient théoriquement se dénaturer plus vite qu'une protéine plus compacte. Cependant, ces mêmes essais suggèrent qu'une réorganisation structurelle se soit produite chez 4H-EV. Considérant ceci, il est fort possible que de nouvelles interactions thermostabilisantes sont présentes chez 4H-EV, autre concept supportant l'idée de changement conformationnel pour ce variant.

Il fut délicat de discerner les sections de pré- et post-transition pour les calculs de dénaturation thermique selon Pace [79]. Il est aussi possible que ces protéines évoluent de manière quasi-continue lors des dénaturations thermiques au lieu de suivre le modèle réversible, coopératif et unimoléculaire.

#### 5.4.4 Dégradations protéolytiques

La résistance aux protéases est un phénomène spécifique. Changer la protéase ou le mélange de protéases utilisé(e)(s) lors d'une expérience peut impliquer une perte de la stabilisation acquise. C'est ce qui se passe dans le cas présent. Lors des expériences d'élimination du site sensible aux protéases chez MB-1TrpHis (projet « NL » présenté dans l'introduction de ce document), les protéases d'*E. coli* exerçaient la pression sélective utilisée. Dans la présente étude, les dégradations protéolytiques ont été effectuées avec le mélange Pronase E. Tout comme elle l'était pour MB-1TrpE<sub>44</sub>V<sub>45</sub>His [58], la présence de la mutation « EV » chez 4H-EV est efficace pour augmenter la résistance aux protéases d'*E. coli*. Il est possible d'affirmer ceci car il y a perte du fragment initial de dégradation (4H) lors de la purification de 4H-EV (voir la figure 4.3). Cependant, cette stabilisation n'est pas retrouvée lorsque les variants sont en présence de Pronase E – la résistance de 4H-EV et de 6H-EV étant même moindre que celle de 4H.

Il est possible que cette diminution résulte d'une reconnaissance directe du site « EV » des mutants par une ou plusieurs des protéases du mélange Pronase E. La dégradation initiale rapide chez les variants « EV » versus 4H peut être expliqué par l'hypothèse suivante. Il est possible que « EV » soit un site pour lequel le mélange de protéases a une haute affinité. Les fragments générés lors de cette dégradation pourraient s'agréger aux protéines intactes, les protégeant de l'attaque des protéases lors de la période plateau (10-30 min). Lors de cette période plateau, les protéases peuvent attaquer des sites pour qui elles ont moins d'affinité. Ceci permettrait la libération des fragments

protecteurs, l'exposition de nouveaux sites de digestion et la suite de la dégradation. Il est intéressant de noter qu'à la fin des incubations, 6H-EV démontre une meilleure résistance à ces protéases que 4H-EV, et ce malgré le fait que 6H-EV présente deux sites « EV ». Si la diminution de résistance provient effectivement de la présence du site « EV », il est envisageable que ces sites soient suffisamment enfoui chez 6H-EV versus chez 4H-EV pour encombrer l'attaque des protéases.

Il est aussi concevable que la diminution de la résistance provienne du fait que ces protéines soient sous une forme complètement différente de celle de la protéine 4H originale. Pour 4H-EV, ce phénomène est connu grâce aux études de fluorescence et de DC. Sous cette nouvelle forme, de nouveaux sites sensibles sont possiblement exposés et atteignables par les protéases du mélange Pronase E. Il est possible que ce soit aussi le cas pour 6H-EV.

## **5.5 Atteintes des objectifs**

La majorité des buts de ce projet ont été atteints. Ce projet a permis

1. de créer les différentes constructions génétiques des divers mutants à l'étude,
2. de produire, purifier et caractériser une partie de ces mutants, et
3. de comparer les propriétés des mutants comportant la modification « EV » (à 4 et 6 hélices) à celles de leur contrepartie originale.

Cependant, un des objectifs visés n'a pas été atteint. Il fut impossible de comparer les propriétés des mutants à 2 et 6 hélices à celles de la protéine originale à 4 hélices (4H) par manque d'accumulation de ces mutants. De plus, il fut impossible de comparer le mutant 2HA avec 2HA-EV pour la même raison. Finalement, la vérification des propriétés de 6H-EV n'a pu être effectuée qu'en partie par manque de pureté des échantillons finaux.

La conclusion finale de ces travaux est donc que la meilleure forme parmi celles explorées pour la séquence MB-1 est un carré à quatre hélices, tel que l'imposait son design initial. De plus, il fut prouvé que la modification N<sub>44</sub>L<sub>45</sub> vers E<sub>44</sub>V<sub>45</sub> aide à l'accumulation des mutants 4H et 6H au sein de la bactérie *E. coli*, mais que la résistance aux protéases serait diminuée en milieu ruminal. Ce dernier fait implique que la modification du site N<sub>44</sub>L<sub>45</sub> chez les divers mutants de MB-1 est inutilisable pour l'objectif d'amélioration de la diète des vaches laitières. Cependant, 4H-EV permettrait peut-être d'atteindre le deuxième but du projet MB-1, soit l'avancement des connaissances fondamentales à propos des carrés d'hélices. Effectivement, il fut suggéré que le changement de conformation protéique lorsque la modification du site NL y est apportée semble stabiliser la structure contre la dénaturation par la chaleur.

## 5.6 Avenues de recherche

Des essais de cristallisation de MB-1 ont déjà été effectués, mais en vain. Des essais sur plusieurs mois n'ont pas fourni de bons cristaux pour permettre la détermination de structure tridimensionnelle par rayons X. Il est possible que la présence de fragment(s) de dénaturation ait nuit à la formation de cristaux [84]. Maintenant que la dégradation est éliminée chez 4H-EV, il serait intéressant de vérifier si la production de cristaux aptes à l'analyse est réalisable. Ceci donnerait une nouvelle information très intéressante pour le projet MB-1, soit la forme exacte de la protéine. Les stratégies d'amélioration de la protéine seraient donc plus ciblées. D'autant plus, elles seraient plus faciles à développer puis à entreprendre.

Bien que 4H-EV soit résistant aux protéases d'*E. coli*, cela ne signifie pas qu'elle soit résistante aux protéases du rumen de la vache laitière. Les tests de dégradation protéolytiques (Pronase E) effectués lors de ce projet le prouvent. Il resterait donc des améliorations à faire. Une détermination de l'emplacement exact du(des) site(s) de dégradation et son remplacement par un(des) site(s)

insensible(s), par exemple, aiderait à la stabilisation de la protéine dans le rumen et sa digestion adéquate chez la vache laitière, but initial du projet MB-1.

Un autre projet pourrait aider à la stabilisation de la protéine. Il existe, au sein de MB-1, un site NL jumeau de celui étudié dans ce projet. Ce deuxième site, le N<sub>93</sub>L<sub>94</sub>, fait partie de l'hélice 4, puis est un des autres sites putatifs de dégradation protéolytique de MB-1TrpHis [58]. Le site N<sub>44</sub>L<sub>45</sub> fut choisi au lieu du site N<sub>93</sub>L<sub>94</sub> car il fut supposé qu'une coupure presque en plein centre de la protéine serait plus nuisible à sa stabilité qu'une coupure en son côté C-terminal. Il serait intéressant de modifier ce site chez 4H et 4H-EV (N<sub>93</sub>L<sub>94</sub>) et chez 6H et 6H-EV (N<sub>142</sub>L<sub>143</sub>). Il serait ensuite possible de vérifier l'impact qu'a cette modification sur la stabilité et la conformation des protéines. Une nouvelle banque de mutants serait construite pour 4H, 4H-EV, 6H et 6H-EV de la même manière que lors des expériences de Sasseville *et al* [58]. D'une part, la modification de ce nouveau site chez 4H pourrait aider à augmenter sa résistance aux protéases d'*E. coli*. Cependant, il est possible que ce changement n'affecte que très peu ou même pas du tout 4H-EV, qui semble déjà immunisé contre l'attaque de ces protéases. D'autre part, l'effet de cette modification pourrait être bénéfique chez 6H et 6H-EV. La modification de N<sub>44</sub>L<sub>45</sub> et de N<sub>93</sub>L<sub>94</sub> chez 6H pour créer 6H-EV aide à l'accumulation de protéine, mais ne suffit pas à la résistance contre les protéases d'*E. coli*. L'ajout de cette nouvelle modification (N<sub>142</sub>L<sub>143</sub>) pourrait aider à favoriser l'accumulation de la structure pleine taille et/ou changer sa structure tertiaire (toujours indéterminée pour ces mutants). Il est également possible que cette modification vienne augmenter la résistance des variants envers les protéases bactériennes et du rumen. La résistance de tous ces mutants contre le mélange Pronase E devra évidemment être investiguée.

## BIBLIOGRAPHIE

1. Church, D.C. (1988) "The Ruminant Animal: Digestive Physiology and Nutrition", Prentice Hall, Ed., Reston U.S.A. pp. 359-360.
2. Chalupa, W. (1975) "Rumen bypass and protection of proteins and amino acids", J. Dairy Sci. 58(8): 1198-1218.
3. Dehority, B.A., Purser, D.B. (1970) "Factors affecting the establishment and numbers of holotrich protozoa in the ovine rumen", J. Anim. Sci. 30(3): 445-449.
4. Miller, W.J. (1979) "Dairy Cattle Feeding and Nutrition", Academic Press, Orlando, U.S.A., pp. 39-49.
5. Stern, M.D., Santos, K.A., Satter, L.D. (1985) "Protein degradation in rumen and amino acid absorption in small intestine of lactating dairy cattle fed heat-treated whole soybeans", J. Dairy Sci. 68(1): 45-56.
6. Ashes, J.R., Gulati, S.K., Scott, T.W. (1995) "The role of rumen protected proteins and energy sources in the diet of ruminants" de "Animal Science Research and Development: Moving Towards a New Century", Ivan M, ed. CFAR Contribution no. 2321, Ministry of Supplies and Services Canada, Ottawa, Canada, pp. 177-185.
7. Rulquin, H., Vérité, R., Guinard, G., Pisulewski, P.M. (1995) "Dairy cows requirement for amino acids" de "Animal Science Research and Development: Moving Toward a New Century", Ivan M, ed. CFAR Contribution no. 2321, Ministry of Supplies and Services Canada, Ottawa, Canada, pp. 143-160.
8. Adisseo (2006) "À propos d'Adisseo – Chiffres clés", Antony Cedex, France, <http://www.an.aventis.com>.
9. Boerjan, W., Bauw, G., Van Montagu, M., Inze, D. (1994) "Distinct phenotypes generated by overexpression and suppression of S-adenosyl-l-methionine synthetase reveal developmental patterns of gene silencing in tobacco", Plant Cell 6(10): 1401-1414.

10. Kim, J., Leustek, T. (2000) "Repression of cystathionine gamma-synthase in *Arabidopsis thaliana* produces partial methionine auxotrophy and developmental abnormalities", *Plant Sci.* 151(1): 9–18.
11. Galili, G. (1995) "Regulation of lysine and threonine synthesis", *Plant Cell* 7(7): 899–906.
12. Karchi, H., Shaul, O., Galili, G. (1993) "Seed specific expression of a bacterial desensitized aspartate kinase increases the production of seed threonine and methionine in transgenic tobacco", *Plant J.* 3(5): 721–727.
13. Falco, S.C., Guida, T., Locke, M., Mauvais, J., Sandres, C., Ward, R.T., Webber, P. (1995) "Transgenic canola and soybean seeds with increased lysine", *Bio/Technology* 13(6): 577–582.
14. Tabe, L.M., Droux, M. (2002) "Limits to sulfur accumulation in transgenic lupin seeds expressing a foreign sulfur-rich protein", *Plant Physiol.* 128(3): 1137–48.
15. Altenbach, S.B, Pearson, K.W, Leung, F.W., Sun, S.S.M. (1987) "Cloning and sequence analysis of a cDNA encoding a Brazil nut protein exceptionally rich in methionine", *Plant Mol. Biol.* 8(3): 239–250.
16. Chui, C.F., Falco, S.C. (1995) "A new methionine-rich seed storage protein from maize", *Plant Physiol.* 107(1): 291.
17. Kortt, A.A., Cladwell, J.B., Lillet, G.G., Higgins, T.J. (1991) "Amino acid and cDNA sequence of a methionine-rich 2S protein from sunflower seed (*Helianthus annuus* L.)", *Eur. J. Biochem.* 195(2): 329–334.
18. Chakraborty, S., Chakraborty, N., Datta, A. (2000) "Increased nutritive value of transgenic potato by expressing a nonallergenic seed albumin gene from *Amaranthus hypochondriacus*", *Proc. Natl. Acad. Sci. USA* 97(7): 3724–3729.
19. Muntz, K., Christov, V., Saalbach, G., Saalbach, I., Waddell, D., Pickardt, T., Schieder, O., Wustenhagen, T. (1998) "Genetic engineering for high methionine grain legumes", *Nahrung* 42(3-4): 125–127.
20. Guenoune, D., Amir, R., Badani, H., Wolf, S., Galili, S. (2002) "Combined expression of S-VSPalpha in two different organelles enhances its



- accumulation and total lysine production in leaves of transgenic tobacco plants", *J. Exp. Bot.* 53(376): 1867-1870.
21. Roesler, K.R., Rao, A.G. (2001) "Rapid gastric fluid digestion and biochemical characterization of engineered proteins enriched in essential amino acids", *J. Agric. Food Chem.* 49(7): 3443-3451.
  22. Forano, E., Flint, H.J. (2000) "Genetically modified organisms: consequences for ruminant health and nutrition", *Ann. Zootech.* 49: 255-271.
  23. Teather, R.M. (1985) "Application of gene manipulation to rumen microflora", *Can. J. Anim. Sci.* 65(3): 563-574.
  24. Jung, R., Falco, S.C. (2000) "Transgenic corn with an improved amino acid composition", papier présenté lors du 8<sup>ième</sup> "International Symposium on Plant Seeds", Gatersleben, Allemagne.
  25. Offredi, F., Dubail, F., Kischel, P., Sarinski, K., Stern, A.S., Van de Weerd, C., Hoch, J.C., Prosperi, C., Francois, J.M., Mayo, S.L., Martial, J.A. (2003) "De novo backbone and sequence design of an idealized alpha/beta-barrel protein: evidence of stable tertiary structure", *J. Mol. Biol.* 325(1): 163-174.
  26. Walsh, S.T., Sukharev, V.I., Betz, S.F., Vekshin, N.L., DeGrado, W.F. (2001) "Hydrophobic core malleability of a de novo designed three-helix bundle protein", *J. Mol. Biol.* 305(2): 361-373.
  27. Johansson, J.S., Gibney, B.R., Skalicky, J.J., Wand, A.J., Dutton, P.L. (1998) "A native-like three- $\alpha$ -helix bundle protein from structure-based redesign: a novel maquette scaffold", *J. Am. Chem. Soc.* 120: 3881-3886.
  28. Schafmeister, C.E., LaPorte, S.L., Miercke, L.J., Stroud, R.M. (1997) "A designed four helix bundle protein with native-like structure", *Nat. Struct. Biol.* 4(12): 1039-1046.
  29. Jiang, X., Bishop, E.J., Farid, R.S.J. (1997) "A de novo designed protein with properties that characterize natural hyperthermophilic proteins", *J. Am. Chem. Soc.* 119(4): 838-839.

30. Brive, L., Dolphin, G.T., Baltzer, L. (1997) "Structure and function of an aromatic ensemble that restricts the dynamics of the hydrophobic core of a designed helix-loop-helix dimer", *J. Am. Chem. Soc.* 119(37): 8598-8607.
31. Gassner, N.C., Baase, W.A., Mooers, B.H., Busam, R.D., Weaver, L.H., Lindstrom, J.D., Quillin, M.L., Matthews, B.W. (2003) "Multiple methionine substitutions are tolerated in T4 lysozyme and have coupled effects on folding and stability", *Biophys. Chem.* 100(1-3): 325-340.
32. Beauregard, M., Hefford, M.A., Teather, R.M. (1994) "Detection of  $\beta$ -galactosidase  $\alpha$ -peptide fusions using a commercially available monoclonal antibody" *Biotechniques* 16(5): 831-834, 836, 838.
33. Jaynes, J., Langridge, P., Anderson, K., Bond, C., Sands, D., Newman, C., Newman, R. (1985) "Construction and expression of sythetic DNA fragments coding for polypeptides with elevated levels of essential amino acids", *Appl. Microbiol. Biotechnol.* 21: 200-205.
34. Beauregard, M., Dupont, C., Teather, R.M., Hefford, M.A. (1995) "Design, expression, and initial characterization of MB1, a de novo protein enriched in essential amino acids", *Biotechnology* 13(9): 974-981.
35. Harbury, P.B., Plecs, J.J., Tidor, B., Alber, T., Kim, P.S. (1998) "High-resolution protein design with backbone freedom", *Science* 282(5395): 1462-1467.
36. Ogihara, N.L., Weiss, M.S., Degrado, W.F., Eisenberg, D. (1997) "The crystal structure of the designed trimeric coiled coil coil-VaLd: implications for engineering crystals and supramolecular assemblies", *Protein Sci.* 6(1): 80-88.
37. Kamtekar, S., Schiffer, J.M., Xiong, H., Babik, J.M., Hecht, M.H. (1993) "Protein design by binary patterning of polar and nonpolar amino acids", *Science* 262(5140): 1680-1685.
38. Pomerantz, J.L., Wolfe, S.A., Pabo, C.O. (1998) "Structure-based design of a dimeric zinc finger protein", *Biochemistry* 37(4): 965-970.
39. Dahiyat, B.I., Sarisky, C.A., Mayo, S.L. (1997) "De novo protein design: towards fully automated sequence selection", *J. Mol. Biol.* 273(4):789-96.

40. Struthers, M.D., Cheng, R.P., Imperiali, B. (1996) "Design of a monomeric 23-residue polypeptide with defined tertiary structure", *Science* 271(5247): 342-345.
41. Sasseville, M. (2004) "Stabilisation de la protéine MB-1Trp par le design de ponts salins à sa surface", Mémoire de maîtrise, Université du Québec à Trois-Rivières.
42. Nishimura, C., Dyson, H.J., Wright, P.E. (2002) "The apomyoglobin folding pathway revisited: structural heterogeneity in the kinetic burst phase intermediate", *J. Mol. Biol.* 322(3):483-9.
43. Morrow, J.A., Hatters, D.M., Lu, B., Hochtl, P., Oberg, K.A., Rupp, B., Weisgraber, K.H. (2002) "Apolipoprotein E4 forms a molten globule. A potential basis for its association with disease", *J. Biol. Chem.* 277(52):50380-5.
44. Permyakov, E.A., Berliner, L.J. (2000) "Alpha-Lactalbumin: structure and function" *FEBS Lett.* 473(3):269-74.
45. Dahiyat, B.I., Mayo, S.L. (1997) "De novo protein design: fully automated sequence selection", *Science* 278(5335):82-7.
46. Chu, R., Takei, J., Knowlton, J.R., Andrykovitch, M., Pei, W., Kajava, A.V., Steinbach, P.J., Ji, X., Bai, Y. (2002) "Redesign of a four-helix bundle protein by phage display coupled with proteolysis and structural characterization by NMR and X-ray crystallography", *J. Mol. Biol.* 323(2):253-62.
47. Wang, L.J., Kong, X.D., Zhang, H.Y., Wang, X.P., Zhang, J. (2000) "Enhancement of the activity of L-aspartase from *Escherichia coli* W by directed evolution", *Biochem, Biophys, Res, Commun.* 276(1):346-9.
48. Betz, S.F., Bryson, J.W., DeGrado, W.F. (1995) "Native-like and structurally characterized designed alpha-helical bundles", *Curr. Opin. Struct. Biol.* 5(4):457-63.
49. Biou, V., Gibrat, J.F., Levin, J.M., Garnier, J. (1988) "Secondary structure prediction: combination of three different methods", *Protein Eng.* 2(3): 185-191.

50. Gagnon, M. (2001) "Création et caractérisation du mutant Trp de la protéine *de novo* MB-1", Mémoire de maîtrise, Université du Québec à Trois-Rivières.
51. Privalov, P.L. (1979) "Stability of proteins: small globular proteins", *Adv. Prot. Chem.* 33: 167-241.
52. MacCallum, J., Hefford, M.A., Omar, S., Beauregard, M. (1997) "Prediction of folding stability and degradability of the *de novo* designed protein MB-1 in cow rumen", *Appl. Biochem. Biotechnol.* 66(1): 83–93.
53. Hefford, M.A., Dupont, C., MacCallum, J., Parker, M.H., Beauregard, M. (1999) "Characterization of MB-1. A dimeric helical protein with a compact core", *European J. Biochem.* 262(2): 467-474.
54. Grundy, J.E., Morrison, J.J., MacCallum, J.D., Wirtanen, L., Beauregard, M. (1998) "Crystallization and stabilization of MB-1, a *de novo* designed protein for optimized feeding technology", *J. Biotechnol.* 63: 9-15.
55. Gagnon, M.C., Williams, M., Doucet, A., Beauregard, M. (2000) "Replacement of tyr62 by trp in the designer protein milk bundle-1 results in significant improvement of conformational stability", *FEBS Lett.* 484(2):144-8.
56. Grundy, J.E., Wirtanen, L.Y., Beauregard, M. (1998) "Addition of a Poly-(6X) His Tag to Milk Bundle-1 and Purification Using Immobilized Metal-Affinity Chromatography", *Prot. Exp. and Pur.* 13(1): 61-66.
57. Beauregard, M., Gagnon, M.C., Doucet, A., Williams, M. (Octobre 2002) "MB-1 Analogues and Uses Thereof", United States Patent and Trademark Office, Continuation-in-Part 10/272,929, demande déposée le 18 octobre 2002.
58. Sasseville, M., St-Louis, C., Khoudi, H., Beauregard, M. (2004) "Controlling proteolytic degradation of the methionine enriched MB-1Trp protein", *Ele. J. Biotech*,  
<http://www.ejbiotechnology.info/content/vol7/issue3/full/10/index.html>.

59. Hill, R.B., Raleigh, D.P., Lombardi, A., DeGrado, W.F. (2000) "*De novo* Design of Helical Bundles as Models for Understanding Protein Folding and Function", *Acc. Chem. Res.* 33(11): 745-754.
60. Myszka, D.G., Chaiken, I.M. (1994) "Design and Characterization of an Intermolecular Antiparallel Coiled Coil Peptide", *Biochemistry* 33: 2363-2372.
61. Hodges, R.S., Saund, A.K., Chong, P.C., St-Pierre, S.A., Reid, R.E. (1981) "Synthetic model for two-stranded alpha-helical coiled-coils. Design, synthesis, and characterization of an 86-residue analog of tropomyosin", *J. Biol. Chem.* 256(3):1214-24.
62. Marti, D.N. Bjelic, S., Lu, M., Bosshard, H.R., Jelesaroy, I. (2004) "Fast Folding of the HIV-1 and SIV gp41 Six-helix Bundles", *J. Mol. Biol.* 336(1):1-8.
63. Ghirlanda, G., Lear, J.D., Ogiwara, N.L., Eisenberg, D., DeGrado, W.F. (2002) "A Hierarchical Approach to the Design of Hexameric Helical Barrels", *J. Mol. Biol.* 319: 243-253.
64. Main, E.R., Xiong, Y., Cocco, M.J., D'Andrea, L., Regan, L. (2003) "Design of stable alpha-helical arrays from an idealized TPR motif", *Structure* 11(5):497-508.
65. Kajander, T., Cortajarena, A.L., Main, E.R., Mochrie, S.G., Regan, L. (2005) "A new folding paradigm for repeat proteins", *J. Am. Chem. Soc.* 127(29):10188-10190.
66. Main, E.R., Scott, K., Jackson, S.E., Regan, L. (2005) "Local and long-range stability in tandemly arrayed tetratricopeptide repeats", *Proc. Natl. Acad. Sci. U.S.A.* 102(16):5721-5726.
67. New England BioLabs (2006), "pMAL<sup>TM</sup> Protein Fusion and Purification System Instruction manual", Ipswich, U.S.A., <http://www.neb.com/nebecomm/ManualFiles/manualE8000.pdf>.
68. Maloy, S. (4 avril 2006) "Expression vectors", San Diego State University, <http://www.sci.sdsu.edu/~smaloy/MicrobialGenetics/topics/in-vitro-genetics/expression-vectors.html>.

69. Stratgene (2006) "QuikChange<sup>®</sup> Site Directed Mutagenesis Kit", La Jolla, U.S.A., <http://www.stratagene.com/products/displayProduct.aspx?pid=131>.
70. Rupp, B. (2005) "CrysPred predictor", q.e.d. life science discoveries, inc, <http://www.ruppweb.org/cryspred/default.html>.
71. Putnam, C. (2005) "Protein Calculator", Scripps Research Institute, <http://www.scripps.edu/~cdputnam/protcalc.html>.
72. Swiss Institute of Bioinformatics, (2005) "Compute pI / Mw tool", ExPASy proteomics server, [http://www.expasy.org/tools/pi\\_tool.html](http://www.expasy.org/tools/pi_tool.html).
73. Qiagen, (Juillet 2002) "QIAquick<sup>®</sup> Spin Handbook for QIAquick PCR purification kit", Mississauga, Ontario, [http://www1.qiagen.com/literature/handbooks/PDF/DNACleanupAndConcentration/QQ\\_Spin/1021422\\_HBQQSpin\\_072002WWW.pdf](http://www1.qiagen.com/literature/handbooks/PDF/DNACleanupAndConcentration/QQ_Spin/1021422_HBQQSpin_072002WWW.pdf).
74. Roche Diagnostics GmbH (2004) "Rapid DNA Ligation Kit", Roche Applied Science, Laval, Quebec, <http://www.rpgroup.caltech.edu/courses/aph162/2006/Protocols/RocheRapidDNALigation.pdf>.
75. Roche Diagnostics GmbH (2004) "The Complete Guide for Protease Inhibition", Roche Applied Science, Laval, Quebec, [http://www.roche-applied-science.com/ProteaseInhibitor/pdf/proteaseinhibition\\_guide.pdf](http://www.roche-applied-science.com/ProteaseInhibitor/pdf/proteaseinhibition_guide.pdf).
76. Millipore Corporation (2004) "Amicon<sup>®</sup> Ultra-15 Centrifugal Filter Devices for volumes up to 15mL – User Guide", Cambridge, Ontario, [http://www.millipore.com/userguides.nsf/dda0cb48c91c0fb6852567430063b5d6/9b4782ed4a3c373285256ff8006f547f/\\$FILE/PR01780.pdf](http://www.millipore.com/userguides.nsf/dda0cb48c91c0fb6852567430063b5d6/9b4782ed4a3c373285256ff8006f547f/$FILE/PR01780.pdf).
77. Sigma-Aldrich (1997) "Protein Determination – Bicinchoninic Acid Method", Oakville, Ontario, [http://www.sigmaaldrich.com/img/assets/18160/Bicinchoninic\\_Acid\\_Method.pdf](http://www.sigmaaldrich.com/img/assets/18160/Bicinchoninic_Acid_Method.pdf).
78. Sreerama, N. (2004) "CDPro – A Software Package for Analysing Protein CD Spectra", Colorado State University, <http://lamar.colostate.edu/~sreeram/CDPro/main.html>.

79. Pace, C.N., Shirley, B.A., Thomson, J.A. (1989) "Measuring the conformational stability of a protein" de "Protein Structure – A Practical Approach", Creighton, T.E., ed., IRL press, Oxford, U.S.A. pp 311-330.
80. Krishnamoorthy, U., Sniffen C.J., Stern, M.D., Van Soest, P.J. (1983) "Evaluation of a mathematical model of rumen digestion and an in vitro simulation of rumen proteolysis to estimate the rumen-undegraded nitrogen content of feedstuffs", Br. J. Nutr. 50(3):555-68.
81. Morrison, J.J., McNally, J.D., Navidzadeh, A. Beauregard, M. (2000) "Development of an optimized feeding technology for dairy cows: improvement in resistance to ruminal proteases in the de novo-designed protein MB-1", Appl. Biochem. Biotechnol. 87(3):247-64.
82. Beauregard, M. (2000) "Méthodes d'études des macromolécules biologiques, BIP-1005, Partie B", Université du Québec à Trois-Rivières, pp. 161-171.
83. Wikipedia contributors (7 Juillet 2006), "X-ray crystallography", Wikipedia, The Free Encyclopedia, [http://en.wikipedia.org/w/index.php?title=X-ray\\_crystallography&oldid=62598428](http://en.wikipedia.org/w/index.php?title=X-ray_crystallography&oldid=62598428).